```
***********************************************************************
   __   __  ______   _____   ______  ______   _     _
  |  \ /  ||  __  | |  ___| |  __  ||  ____| | |   | |
  |   V   || |__| | | |___  | |__| || |      | |___| |
  | |\_/| ||  ____| |___  | |  __ / | |      |  ___  |
  | |   | || |       ___| | | |\ \  | |____  | |   | |
  |_|   |_||_|      |_____| |_| \_\ |______| |_|   |_|  (TM)
***********************************************************************

         Release 2 1D John F. Collins, Biocomputing Research Unit
         Copyright (c) 1993, 1994, 1995 University of Edinburgh, U K
                  Distribution rights by IntelliGenetics, Inc

MPsrch_pp   protein - protein database search, using Smith-Waterman algorithm

Run on:          Thu Dec 18 09:35:29 1997;  MasPar time 6.92 Seconds
                                            724.336 Million cell updates/sec
Tabular output not generated

Title:           >US-08-826-361A-4
Description:     (1-233) from US08826361A.pep
Perfect Score:   1726
Sequence:        1 LVLTLLEAEPPHVLISRPSA..........CKNVVPVYDLLLEMLNAHVL 233

Scoring table:   PAM 150
                 Gap 11

Searched:        59021 seqs, 21210398 residues

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database:        swiss-prot34
                   1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
                   8:part8 9:part9 10:part10 11:part11

Statistics:      Mean 47.263;  Variance 91.493;  scale 0.517

       Pred. No. is the number of results predicted by chance to have a
       score greater than or equal to the score of the result being printed,
       and is derived by analysis of the total score distribution.

                                 SUMMARIES
                  %
Result          Query
  No.   Score   Match Length DB  ID          Description             Pred. No.
-------------------------------------------------------------------------------
    1    1142    66.2    589  3  ESTR_CHICK  ESTROGEN RECEPTOR (ER  3.01e-210
    2    1126    65.2    595  3  ESTR_HUMAN  ESTROGEN RECEPTOR (ER  8.42e-207
    3    1122    65.0    599  3  ESTR_MOUSE  ESTROGEN RECEPTOR (ER  5.12e-206
    4    1118    64.8    600  3  ESTR_RAT    ESTROGEN RECEPTOR (ER  4.45e-205
    5    1084    62.8    575  3  ESTR_ORYLA  ESTROGEN RECEPTOR (ER  9.17e-198
    6    1070    62.0    314  3  ESTR_OREAU  ESTROGEN RECEPTOR (ER  9.36e-195
    7    1066    61.8    535  3  ESTR_SALSA  ESTROGEN RECEPTOR (ER  6.77e-194
    8    1048    60.7    574  3  ESTR_SALIR  ESTROGEN RECEPTOR (ER  4.97e-190
    9     579    33.5    133  3  ESTR_BOVIN  ESTROGEN RECEPTOR (ER  7.90e-91
   10     499    28.9    121  3  ESTR_MACMU  ESTROGEN RECEPTOR (ER  2.19e-74
   11     486    28.2    433  3  ERR2_HUMAN  STEROID HORMONE RECEP  3.73e-71
   12     444    25.7    521  3  ERR1_HUMAN  STEROID HORMONE RECEP  3.13e-63
   13     338    19.6    933  7  PRGR_HUMAN  PROGESTERONE RECEPTOR  3.30e-42
   14     336    19.5    923  7  PRGR_MOUSE  PROGESTERONE RECEPTOR  8.09e-42
   15     336    19.5    930  7  PRGR_RABIT  PROGESTERONE RECEPTOR  8.08e-42
   16     322    18.7    758  4  GCR_ONCMY   GLUCOCORTICOID RECEPT  4.15e-39
   17     314    18.2    776  4  GCR_XENLA   GLUCOCORTICOID RECEPT  1.44e-37
   18     311    18.0    783  4  GCR_MOUSE   GLUCOCORTICOID RECEPT  5.43e-37
   19     307    17.8    777  4  GCR_HUMAN   GLUCOCORTICOID RECEPT  3.17e-36
   20     304    17.6    771  4  GCR_CAVPO   GLUCOCORTICOID RECEPT  1.19e-35
   21     299    17.3    533  8  RRXB_HUMAN  RETINOIC ACID RECEPTO  1.07e-34
   22     296    17.1    458  8  RRXB_RAT    RETINOIC ACID RECEPTO  3.99e-34
   23     296    17.1    520  8  RRXB_MOUSE  RETINOIC ACID RECEPTO  3.99e-34
   24     293    17.0    795  4  GCR_RAT     GLUCOCORTICOID RECEPT  1.48e-33
   25     286    16.6    462  8  RRXA_HUMAN  RETINOIC ACID RECEPTO  3.15e-32
   26     286    16.6    467  8  RRXA_CHICK  RETINOIC ACID RECEPTO  3.15e-32
   27     286    16.6    467  8  RRXA_RAT    RETINOIC ACID RECEPTO  3.15e-32
   28     286    16.6    467  8  RRXA_MOUSE  RETINOIC ACID RECEPTO  3.15e-32
   29     285    16.5    488  8  RRXA_XENLA  RETINOIC ACID RECEPTO  2.04e-32
   30     285    16.5    786  7  PRGR_CHICK  PROGESTERONE RECEPTOR  2.04e-32
   31     285    16.5    984  6  MCR_HUMAN   MINERALOCORTICOID REC  2.04e-32
   32     284    16.5    981  6  MCR_RAT     MINERALOCORTICOID REC  7.51e-32
   33     283    16.4    463  8  RRXG_MOUSE  RETINOIC ACID RECEPTO  1.16e-31
   34     282    16.3    463  8  RRXG_HUMAN  RETINOIC ACID RECEPTO  1.79e-31
   35     274    15.9    470  8  RXRG_XENLA  RETINOIC ACID RECEPTO  5.72e-30
   36     264    15.3    793  1  ANDR_RABIT  ANDROGEN RECEPTOR (FR  4.23e-28
   37     264    15.3    899  1  ANDR_MOUSE  ANDROGEN RECEPTOR.     4.23e-28
   38     264    15.3    902  1  ANDR_RAT    ANDROGEN RECEPTOR.     4.23e-28
   39     264    15.3    919  1  ANDR_HUMAN  ANDROGEN RECEPTOR.     4.23e-28
   40     233    13.5    560  6  LRH1_MOUSE  LIVER RECEPTOR HOMOLO  2.10e-22
   41     226    13.1    597  6  NGFI_RAT    NERVE GROWTH FACTOR I  3.86e-21
   42     226    13.1    598  6  NGFI_HUMAN  EARLY RESPONSE PROTEI  3.86e-21
   43     226    13.1    598  6  NGFI_CANFA  ORPHAN NUCLEAR RECEPT  3.86e-21
   44     216    12.5    601  6  NGFI_MOUSE  NUCLEAR HORMONE RECEP  2.37e-19
   45     212    12.3    808  4  FTFB_DROME  NUCLEAR HORMONE RECEP  1.21e-18

                                 ALIGNMENTS

RESULT     1
ID   ESTR_CHICK     STANDARD;      PRT;   589 AA.
AC   P06212;
DT   01-JAN-1988 (REL. 06, CREATED)
DT   01-JAN-1988 (REL. 06, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   ESTROGEN RECEPTOR (ER).
GN   ESR.
OS   GALLUS GALLUS (CHICKEN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;
OC   GALLIFORMES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 86247578.
RA   KRUST A., GREEN S., ARGOS P., KUMAR V., WALTER P., BORNERT J.-M.,
RA   CHAMBON P.;
RL   EMBO J. 5:891-897(1986).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- IN THE ABSENCE OF LIGAND, STEROID HORMONE RECEPTORS ARE THOUGHT
CC       TO BE WEAKLY ASSOCIATED WITH NUCLEAR COMPONENTS; HORMONE BINDING
CC       GREATLY INCREASES RECEPTOR AFFINITY. THE HORMONE-RECEPTOR
CC       COMPLEX APPEARS TO RECOGNIZE DISCRETE DNA SEQUENCES UPSTREAM OF
CC       TRANSCRIPTIONAL START SITES.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; X03805; G63380; -.
DR   PIR; S07192; S07192.
DR   HSSP; P06536; 1GDC.
DR   PROSITE; PS00031; STEROID_FINGER.
DR   TRANSFAC; T00264; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
FT   DOMAIN          1    178       MODULATING.
FT   DNA_BIND      179    244       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING       179    199       C4-TYPE.
FT   ZN_FING       215    239       C4-TYPE.
FT   DOMAIN        245    304       HINGE.
FT   DOMAIN        305    589       STEROID-BINDING.
SQ   SEQUENCE   589 AA;  66746 MW;  85FF1F4A CRC32;
```

```
  Query Match             65.2%;  Score 1142;  DB 3;  Length 599;
  Best Local Similarity   62.1%;  Pred. No. 3.01e-210;
  Matches  146;  Conservative   53;  Mismatches  34;  Indels   2;  Gaps   2;

Db    309 mvsalleaeppivyseydpnrpfneasmmtlltnladrelvhminwakrvpgfvdltlhd 368
          :| :||||||| |      |: || |||||  ||:|||:||||||:|||::|||||:|:| |
Qy      1 LVLTLLEAEPPHVLIS-RPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 59

Db    369 qvhllecawleilmiglvwrsmehpgkllfapnlllldrnqgkcvegmveifdmllataar 428
          ||:|||  |:|:||:||:|||::|||||:|||:|:|||::||||||::||||||||||::|
Qy     60 QVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSR 119

Db    429 frmmnlqgeefvclksiillnsgvytflsstlksleerdyihrvldkitdtlihlmaksg 488
          || ::||  |::|:|::||||::| ::::|   :    |    | :|: :||:|: ::||||
Qy    120 FRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAH-LLNAVTDALVWVIAKSG 178

Db    489 lslqqqhrrlaqllliilshirhmsnkgmehlynmkcknvvplydlllemldahrl 543
          :| |||  ||| ||::|||:|| |||||||| ||||||||||:||||||||:|| |
Qy    179 ISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNAHVL 233


RESULT     2
ID   ESTR_HUMAN     STANDARD;      PRT;   595 AA
AC   P03372;
DT   21-JUL-1986 (REL. 01, CREATED)
DT   01-JUN 1994 (REL. 29, LAST SEQUENCE UPDATE)
DT   01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE   ESTROGEN RECEPTOR (ER).
GN   ESR.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 86122927.
RA   GREENE G.L., GILNA P., WATERFIELD M., BAKER A., HORT Y., SHINE J.;
RL   SCIENCE 231:1150-1154(1986)
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 86146892.
RA   GREEN S., WALTER P., KUMAR V., KRUST A., BORNERT J.-M., ARGOS P.,
RA   CHAMBON P.;
RL   NATURE 320:134-139(1986).
RN   [3]
RP   STRUCTURE BY NMR OF 180-262.
RX   MEDLINE; 91061891.
RA   SCHWABE J.W.R., NEUHAUS D., RHODES D.;
RL   NATURE 348:458-461(1990).
RN   [4]
RP   VARIANT VAL-400.
RX   MEDLINE; 90065402
RA   TORA L., MULLICK A., METZGER D., PONGLIKITMONGKOL M., PARK I.,
RA   CHAMBON P.;
RL   EMBO J. 8:1981-1986(1989).
RN   [5]
RP   VARIANT ALA-447.
RX   MEDLINE; 92250650.
RA   REESE J.C., KATZENELLENBOGEN B.S.;
RL   J. BIOL. CHEM. 267:9868-9873(1992).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN
CC   -!  IN THE ABSENCE OF LIGAND, STEROID HORMONE RECEPTORS ARE THOUGHT
CC       TO BE WEAKLY ASSOCIATED WITH NUCLEAR COMPONENTS; HORMONE BINDING
CC       GREATLY INCREASES RECEPTOR AFFINITY. THE HORMONE-RECEPTOR
CC       COMPLEX APPEARS TO RECOGNIZE DISCRETE DNA SEQUENCES UPSTREAM OF
CC       TRANSCRIPTIONAL START SITES.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE SIPE II,TBYE II,RETINOI
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; X03635; G31234; -.
DR   EMBL; M12674; G182193; -.
DR   PIR; A03244; QRHUE.
DR   HSSP; P06536; 1GDC.
DR   MIM; 133430; -.
DR   PROSITE; PS00031; STEROID_FINGER.
DR   TRANSFAC; T00261; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER, STEROID-BINDING, POLYMORPHISM.
FT   DOMAIN        1     184       MODULATING.
FT   DNA_BIND    185     250       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     185     205       C4-TYPE.
FT   ZN_FING     221     245       C4-TYPE.
FT   DOMAIN      251     310       HINGE.
FT   DOMAIN      311     551       STEROID-BINDING.
FT   VARIANT     400     400       G -> V (DESTABILIZES THE RECEPTOR AND
FT                                 DECREASES ITS AFFINITY FOR ESTRADIOL AT
FT                                 25 DEGREES CELSIUS, BUT NOT AT 4 DEGREES
FT                                 CELSIUS).
FT   VARIANT     447     447       C -> A (LOSS OF HORMONE BINDING CAPACITY
FT                                 AND TEMPERATURE-SENSITIVE LOSS IN DNA-
FT                                 BINDING).
SQ   SEQUENCE   595 AA;  66216 MW;  D6945BB9 CRC32;

  Query Match             65.2%;  Score 1126;  DB 3;  Length 595;
  Best Local Similarity   60.2%;  Pred. No. 8.42e-207;
  Matches  142,  Conservative   61,  Mismatches  29;  Indels   4;  Gaps   3;

Db    315 mvsalldaepp-ilyseydptrpfseasmmglltnladrelvhminwakrvpgfvdltlh 373
          :| :||:|||| :| |    |: ||:|||||  ||:|||:||||||:|||::|||||:|:|
Qy      1 LVLTLLEAEPPHVLISR-PSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLF 58

Db    374 dqvhllecawleilmiglvwrsmehpgkllfapnlllldrnqgkcvegmveifdmllatss 433
          |||:|||  |:|:||:||:|||::|||||:|||:|:|||::||||||::||||||||||:|
Qy     59 DQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTS 118

Db    434 rfrmmnlqgeefvclksiillnsgvytflsstlksleekdhihrvldkitdtlihlmaka 493
          ||| ::||  |::|:|::||||::| ::::|   : :      : ::|: :||:|: ::||:
Qy    119 RFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSRKLAHLLNAVTDALVWVIAKS 177

Db    494 gltlqqqhqrlaqlllilshirhmsnkgmehlysmkcknvvplydlllemldahrl 549
          |:: |||  ||| ||::|||:|| |||||||| :||||||||:||||||||:|| |
Qy    178 GISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNAHVL 233


RESULT     3
ID   ESTR_MOUSE     STANDARD;      PRT;   599 AA.
AC   P19785;
DT   01-FEB-1991 (REL. 17, CREATED)
DT   01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
DT   01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE   ESTROGEN RECEPTOR (ER).
GN   ESR OR ESTR.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=UTERUS;
RX   MEDLINE; 91042558.
RA   WHITE R., LEES J.A., NEEDHAM M., HAM J., PARKER M.;
RL   MOL. ENDOCRINOL. 1:735-744(1987).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- IN THE ABSENCE OF LIGAND, STEROID HORMONE RECEPTORS ARE THOUGHT
CC       TO BE WEAKLY ASSOCIATED WITH NUCLEAR COMPONENTS. HORMONE BINDING
CC       GREATLY INCREASES RECEPTOR AFFINITY. THE HORMONE-RECEPTOR
CC       COMPLEX APPEARS TO RECOGNIZE DISCRETE DNA SEQUENCES UPSTREAM OF
CC       TRANSCRIPTIONAL START SITES.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; M38651; G193180; -.
DR   PIR; A40061; A40061.
DR   HSSP; P06536; 1GDC.
DR   PROSITE; PS00031; STEROID_FINGER.
DR   TRANSFAC; T00259; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
FT   DOMAIN        1    188       MODULATING.
FT   DNA_BIND    189    254       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     189    209       C4-TYPE.
FT   ZN_FING     225    249       C4-TYPE.
FT   DOMAIN      255    599       STEROID-BINDING.
FT   DOMAIN       64     72       POLY-ALA.
SQ   SEQUENCE   599 AA;  66455 MW;  FE09EADB CRC32;

  Query Match            65.0%;  Score 1122;  DB 3;  Length 599;
  Best Local Similarity  60.0%;  Pred. No. 6.12e-206;
  Matches  141;  Conservative   59;  Mismatches   33;  Indels    2;  Gaps    2;

Db    319 mvsalldaeppmiyseydpsrpfseasmmglltnladrelvhminwakrvpgfgdlnlhd 378
          :| :||:||||  :      || ||:|||||  ||:||:|||||||:|||::||| :|:| |
Qy      1 LVLTLLEAEPPHVLIS-RPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 59

Db    379 qvhllecawleilmiglvwrsmehpgkllfapnllldrnqgkcvegmveifdmllatssr 438
          ||:|||  |:|:||:|| |||  ||||| |||  :  |||   ||||||  |||||||||:||
Qy     60 QVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSR 119

Db    439 frmmnlqgeefvclksiillnsgvytflsstlksleekdhihrvldkitdtlihlmakag 498
          :|  ::||  |::|:|::||||||::| :::  |       |   ||  |       || |
Qy    120 FRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSRKLAHLLNAVTDALVWVIAKSG 178

Db    499 ltlqqqhrrlaqllllilshirhmsnkgmehlynmkcknvvplydlllemldahrl 553
          :: |||  ||| ||::|||:|| |||||||| |||||||||:|||||||||:|| |
Qy    179 ISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNAHVL 233


RESULT     4
ID   ESTR_RAT       STANDARD;      PRT;   600 AA.
AC   P06211;
DT   01-JAN-1988 (REL. 06, CREATED)
DT   01-JAN-1988 (REL. 06, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   ESTROGEN RECEPTOR (ER).
GN   ESR OR ESTR.
OS   RATTUS NORVEGICUS (RAT).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WISTAR;
RA   MURAMATSU M.;
RL   SUBMITTED (MAR-1987) TO EMBL/GENBANK/DDBJ DATA BANKS
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 87174780.
RA   KOIKE S., SAKAI M.;
RL   NUCLEIC ACIDS RES. 15:2499-2513(1987).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY, TISSUE=UTERUS.
RA   MAGGI A.M.A.;
RL   SUBMITTED (JUN-1991) TO EMBL/GENBANK/DDBJ DATA BANKS
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES
CC   -!- SUBCELLULAR LOCATION: NUCLEAR
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN
CC   -!- IN THE ABSENCE OF LIGAND, STEROID HORMONE RECEPTORS ARE THOUGHT
CC       TO BE WEAKLY ASSOCIATED WITH NUCLEAR COMPONENTS; HORMONE BINDING
```

```
CC       GREATLY INCREASES RECEPTOR AFFINITY. THE HORMONE-RECEPTOR
CC       COMPLEX APPEARS TO RECOGNIZE DISCRETE DNA SEQUENCES UPSTREAM OF
CC       TRANSCRIPTIONAL START SITES.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; Y00102; G56111; -.
DR   EMBL; X61098; G56121; -.
DR   PIR; S07379; S07379.
DR   HSSP; P06536; 1GDC.
DR   PROSITE; PS00031; STEROID_FINGER.
DR   TRANSFAC; T00258; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
FT   DOMAIN        1    189       MODULATING.
FT   DNA_BIND    190    255       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     190    210       C4-TYPE.
FT   ZN_FING     226    250       C4-TYPE.
FT   DOMAIN      256    600       STEROID-BINDING.
FT   DOMAIN       64     71       POLY-ALA.
FT   CONFLICT    468    468       N -> I (IN REF. 3).
SQ   SEQUENCE   600 AA;  67030 MW;  400C59B2 CRC32;

  Query Match            64.8%;  Score 1118;  DB 3;  Length 600;
  Best Local Similarity  59.6%;  Pred. No. 4.45e-205;
  Matches  140;  Conservative   59;  Mismatches   34;  Indels    2;  Gaps    2;

Db    320 mvsalldaepplyseydpsrpfseasmmglltnladrelvhminwakrvpgfgdlnlhd 379
          :| :||:||||  :      || ||:|||||  ||:|||:||||||:|||::||| :|:| |
Qy      1 LVLTLLEAEPPHVLIS-RPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 59

Db    380 qvhllecawleilmiglvwrsmehpgkllfapnllldrnqgkcvegmveifdmllatssr 439
          ||:|||  |:|:||:|||:||  ||||||:|||:|:||||:|||||||::||||||||:||
Qy     60 QVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSR 119

Db    440 frmmnlqgeefvclksiillnsgvytflsstlksleekdhihrvldkindtlihlmakag 499
          ||  ::||  |::|:|::|||||::| ::::|   : :     : ::| : |:|: ::|:|
Qy    120 FRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSRKLAHLLNAVTDALVWVIAKSG 178

Db    500 ltlqqqhrrlaqllllilshirhmsnkgmehlynmkcknvvplydlllemldahrl 554
          :: |||  ||| ||::|||:|| |||||||| ||||||||||:||||||||:||
Qy    179 ISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNAHVL 233


RESULT     5
ID   ESTR_ORYLA     STANDARD;      PRT;   575 AA.
AC   P50241;
DT   01-OCT-1996 (REL. 34, CREATED)
DT   01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   ESTROGEN RECEPTOR (ER).
GN   ESR.
OS   ORYZIAS LATIPES (MEDAKA FISH).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; PISCES; GNATHOSTOMATA;
OC   OSTEICHTHYES; ACTINOPTERYGII; ATHERINIFORMES
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RA   OKADA H.;
RL   SUBMITTED (APR-1994) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; D28954; G602326; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
FT   DOMAIN        1    140       MODULATING.
FT   DNA_BIND    141    206       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     141    161       C4-TYPE.
FT   ZN_FING     177    201       C4-TYPE.
FT   DOMAIN      207    575       STEROID-BINDING.
FT   DOMAIN      254    262       POLY-GLY.
FT   DOMAIN      275    278       POLY-LEU.
FT   DOMAIN      466    470       POLY-LEU.
FT   DOMAIN      531    534       POLY-GLY.
SQ   SEQUENCE   575 AA;  62987 MW;  D9FF40D0 CRC32;

  Query Match             62.8%;  Score 1084;  DB 3;  Length 575;
  Best Local Similarity   61.3%;  Pred. No. 9.17e-198;
  Matches  144;  Conservative   45;  Mismatches  43;  Indels    3;  Gaps    3;

Db     274 vlllqqaeppilcsrqklsrpytevtmmtlltsmadkelvhmiawakklpgflqlslhd 333
           || ||::  | :| ||    : |:|| :||  || :|||||||||:||||:|||::||| |
Qy       2 VLTLLEAEPPHVLISRPS-A-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 59

Db     334 qvlllesswlevlmigliwrsihcpgklifaqdlildrneqdcveqmteifdmllatasr 393
           |. |||| |:||||:||:||||  ||||||| ||:|||:|| ||||: |||||||||:||
Qy      60 QVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSR 119

Db     394 frvlklkpeefvclkaiillnsgafsfctgtmeplhnsaavqsmldtitdalihyisqsg 453
           || |||    |::|:||:|||||: ::: |:| ::  :|  :  :|:::||||:  |: ||
Qy     120 FRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSRKLAHLLNAVTDALVWVIAKSG 178

Db     454 ylaqeqarrqaqllllshirhmsnkgmehlysmkcknkvplydlllemldahrl 508
              :|:|: | | ||:||||:|| |||||||| :||||| ||:|||||||||:|| |
Qy     179 ISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNAHVL 233


RESULT    6
ID   ESTR_OREAU      STANDARD;      PRT;   314 AA.
AC   P50240;
DT   01-OCT-1996 (REL. 34, CREATED)
DT   01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   ESTROGEN RECEPTOR (ER) (FRAGMENT).
GN   ESR.
OS   OREOCHROMIS AUREUS.
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; PISCES; GNATHOSTOMATA;
OC   OSTEICHTHYES; ACTINOPTERYGII; PERCIFORMES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BLOOD;
RX   MEDLINE; 96272302.
RA   TAN N., LAM T., DING J.;
RL   DNA SEQ. 5:359-370(1995).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; Z46665; E212276; -.
DR   EMBL; Z46666; E212276; JOINED.
DR   EMBL; Z46667; E212276; JOINED.
DR   EMBL; Z46668; E212276; JOINED.
DR   EMBL; Z46669; E212276; JOINED.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
FT   NON_TER       1      1
FT   DOMAIN       <1    314       STEROID-BINDING (POTENTIAL).
SQ   SEQUENCE   314 AA;  34872 MW;  0B18C899 CRC32;

  Query Match             62.0%;  Score 1070;  DB 3;  Length 314;
  Best Local Similarity   61.4%;  Pred. No. 9.36e-195;
  Matches  143;  Conservative   43;  Mismatches  43;  Indels    4;  Gaps    4;

Db       1 vlllqqaeppilcsrqkmsrpytevtimtlltsmadkelvhmitwakklpgflqlslhd 60
           || ||::  | :| ||    : |:|| ::|  || :|||||||||:||||:|||::||| |
Qy       2 VLTLLEAEPPHVLISRPS-A-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 59

Db      61 qvlllesswlevlmigllwrsiqcpgklifeedlildrneqtcveqmaeifdmllatv-r 119
           || |||| |.||||.||:||||: ||||||  ||:|||:|| ||||: ||||||||   :
Qy      60 QVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSR 119

Db     120 frvlklkpeefvclkaiillnsgafsfctgtmeplhdsvavqhmldtitdalifhishfq 179
           || |||    |::|:||:|||||: ::: |:| ::    |   : |:|:::||||:  |:  |
Qy     120 FRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSRKLAHLLNAVTDALVWVIAKSG 178

Db     180 csaqqqsrrqaqllllshirhmsnkgmehlysmkcknkvplydlllemldah 232
             |:|||| | | ||:||||:|| |||||||| :||||| ||:|||||||||:||
Qy     179 ISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNAH 231


RESULT    7
ID   ESTR_SALSA      STANDARD;      PRT;   535 AA.
AC   P50242;
DT   01-OCT-1996 (REL. 34, CREATED)
DT   01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   ESTROGEN RECEPTOR (ER) (FRAGMENT).
GN   ESR.
OS   SALMO SALAR (ATLANTIC SALMON).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; PISCES; GNATHOSTOMATA;
OC   OSTEICHTHYES; ACTINOPTERYGII; SALMONIFORMES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RA   ROGERS S.A., LLEWELLYN L., RAMSURN V.P., SWEENEY G.E., WIGHAM T.;
RL   SUBMITTED (JUL-1995) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; X89959; G929902; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
FT   NON_TER       1      1
FT   DOMAIN       <1    104       MODULATING (POTENTIAL).
FT   DNA_BIND    105    170       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     105    125       C4-TYPE.
FT   ZN_FING     141    165       C4-TYPE.
FT   DOMAIN      171    535       STEROID-BINDING (POTENTIAL).
SQ   SEQUENCE   535 AA;  58900 MW;  0F71CD0C CRC32;

  Query Match             61.8%;  Score 1066;  DB 3;  Length 535;
  Best Local Similarity   60.9%;  Pred. No. 6.77e-194;
  Matches  143;  Conservative   40;  Mismatches  49;  Indels    3;  Gaps    3;

Db     241 vlfllqqaeppalcsrqkvarpytevtmmtlltsmadkelvhmlawakkvpgfqelslhd 300
           || ||::  |  | ||    | |:|| :||  || :|||||||||:||||:||| |||| |
Qy       2 VLTLLEAEPPHVLISRPS-A-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 59

Db     301 qvqllesswlevlmigliwrsihcpgklifaqdlildrseqdcveqmaeifdmllatvsr 360
           ||.|||| |:||||:||:||||  ||||||| ||:||| || ||||: ||||||||| ||
Qy      60 QVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSR 119

Db     361 frmlklkpeefvclkaiillnsgafsfcsnsveslhnssavesmldnitdalihhishsg 420
           || |||    |::|:||:|||||: ::: ::  :   :|   :  :|: :||||:  |: ||
Qy     120 FRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSRKLAHLLNAVTDALVWVIAKSG 178

Db     421 asvqqqprrqvqllllshirhmsnkgmehlysikcknkvplydlllemldahrl 475
             | |||: |   ||:||||:|| |||||||| ::|||| ||:||||||||::| |
Qy     179 ISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNAHVL 233
```

```
RESULT     8
ID   ESTR_SALIR     STANDARD;      PRT;   574 AA.
AC   P16058;
DT   01-APR-1990 (REL. 14, CREATED)
DT   01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   ESTROGEN RECEPTOR (ER).
GN   ESR.
OS   SALMO IRIDEUS (RAINBOW TROUT).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; PISCES; GNATHOSTOMATA;
OC   OSTEICHTHYES; ACTINOPTERYGII; SALMONIFORMES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 91006824.
RA   PAKDEL F., LE GAC F., LE GOFF P., VALOTAIRE Y.;
RL   MOL. CELL. ENDOCRINOL. 71:195-204(1990).
RN   [2]
RP   SEQUENCE OF 150-574 FROM N.A.
RC   TISSUE=LIVER;
RX   MEDLINE; 89127284.
RA   PAKDEL F., LE GUELLEC C., VAILLANT C., LE ROUX M.G., VALOTAIRE Y.;
RL   MOL. ENDOCRINOL. 3:44-51(1989).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- IN THE ABSENCE OF LIGAND, STEROID HORMONE RECEPTORS ARE THOUGHT
CC       TO BE WEAKLY ASSOCIATED WITH NUCLEAR COMPONENTS; HORMONE BINDING
CC       GREATLY INCREASES RECEPTOR AFFINITY. THE HORMONE-RECEPTOR
CC       COMPLEX APPEARS TO RECOGNIZE DISCRETE DNA SEQUENCES UPSTREAM OF
CC       TRANSCRIPTIONAL START SITES.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; M31559; G213784; -.
DR   PIR; A37197; A37197.
DR   PIR; A40070; A40070.
DR   HSSP; P06536; 1GDC.
DR   PROSITE; PS00031; STEROID_FINGER.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
FT   DOMAIN        1    146       MODULATING (POTENTIAL).
FT   DNA_BIND    147    212       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     147    167       C4-TYPE.
FT   ZN_FING     183    207       C4-TYPE.
FT   DOMAIN      213    574       STEROID-BINDING (POTENTIAL).
SQ   SEQUENCE   574 AA;  62872 MW;  B391ADC2 CRC32;

  Query Match             60.7%;  Score 1048;  DB 3;  Length 574;
  Best Local Similarity  60.4%;  Pred. No. 4.97e-190;
  Matches  142;  Conservative   42;  Mismatches  47;  Indels   4;  Gaps   4;

Db     291 vlfllqqtp-alvsrqkvarpytevtmmtlltsmadkelvhmiawakkvpgfqelslhd 339
           || ||:::: |  | ||    | |:|| :||  || :|||||||||:|||||:||| |||| |
Qy       2 VLTLLEAEPPHVLISRPS-A-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 59

Db     340 qvqllesswlevlmiglmwrsihcpgklifaqdlildrsegdcvegmaeifdmllatvsr 399
           ||:|||| |:||||:||:||||  ||||||| ||:||| || ||||: |||||||| ||
Qy      60 QVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSR 119

Db     400 fqmlklkpeefvclkaiillnpgafsfcsnsveslhnssavesmldnitdalihhishsg 459
           |  |||   |:::|::||||:: ::: : : ..  .|  .  .|. .||||  |  ||
Qy     120 FRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSRKLAHLLNAVTDALVWVIAKSG 178

Db     460 asvqqqprrqaqlllllshirhmsnkgmehlysikckakvplydlllemldghrl 514
            | |||: | | ||:||||:|| |||||||| ..|||| ||.|||||||||..| |
Qy     179 ISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNAHVL 233


RESULT     9
ID   ESTR_BOVIN     STANDARD;      PRT;   133 AA.
AC   P49884;
DT   01-OCT-1996 (REL. 34, CREATED)
DT   01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   ESTROGEN RECEPTOR (ER) (FRAGMENT).
GN   ESR.
OS   BOS TAURUS (BOVINE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; ARTIODACTYLA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SIMMENTAL BREED; TISSUE=UTERUS;
RA   HAGEN-MANN K., MANN W., MEYER H.H.D.;
RL   SUBMITTED (MAY 1992) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; X66841; G335; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
FT   NON_TER       1      1
FT   DOMAIN       <1   >133       STEROID-BINDING.
FT   NON_TER     133    133
SQ   SEQUENCE   133 AA;  15298 MW;  AE22720E CRC32;

  Query Match             35.5%;  Score 573;  DB 3;  Length 133;
  Best Local Similarity  52.7%;  Pred. No. 7.90e-91;
  Matches   69;  Conservative   43;  Mismatches  18;  Indels   1;  Gaps   1;

Db       1 ltlhdqvhllecawleilmiglvwrsmehpvkllfapnllldrnqgkcvegmveifdmll 60
           |.| |||.|||  |:|:||:||:|||::|| ||:|||::|:|||::||||::|||'|:|
Qy      55 LSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLL 114

Db      61 atssrfrmmnlqgeefvclksiillnsgvytflsstlksleekdhihrvldkitdtlihl 120
           ||.|||| ..||  |..|.|..||||||::| ::::|  : :    : ::|: :|.:.: :
Qy     115 AIISRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSRKLAHLLNAVTDALVWV 173

Db     121 makagltlqqq 131
           :||:||: |||
Qy     174 IAKSGISSQQQ 184


RESULT    10
ID   ESTR_MACMU     STANDARD;      PRT;   121 AA.
AC   P49886;
DT   01-OCT-1996 (REL. 34, CREATED)
DT   01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   ESTROGEN RECEPTOR (ER) (FRAGMENT).
GN   ESR.
OS   MACACA MULATTA (RHESUS MACAQUE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=OVARY;
RX   MEDLINE; 94283272.
RA   CHANDRASEKHER Y.A., MELNER M.H., NAGALLA S.R., STOUFFER R.L.;
RL   ENDOCRINOLOGY 135:307-314(1994).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC         NUCLEAR HORMONE RECEPTORS.
DR    EMBL; S71040; G547183; -.
KW    RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW    ZINC-FINGER; STEROID-BINDING.
FT    NON_TER         1       1
FT    DOMAIN         <1    >121        STEROID BINDING.
FT    NON_TER       121     121
SQ    SEQUENCE   121 AA;  13836 MW;  E59BD9E0 CRC32;

  Query Match             28.9%;  Score 499;  DB 3;  Length 121;
  Best Local Similarity   52.1%;  Pred. No. 2.18e-74;
  Matches    63;  Conservative   38;  Mismatches  19;  Indels   1;  Gaps   1;

Db      1 lfapnllldrnqgkcveqmvesfdmllatssrfrmmnlqgeefvclksiillnsgvytfl 60
          :|||:|:|||::||||||::| |||||||:|||| ::||  |:.:.:..|||||.:| ..
Qy     88 IFAPDLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLV 147

Db     61 sstlksleekdhihrvldkitdtlihlmakagltlqqqhrrlaqlllilshirhmsnkgm 120
          ::|  : :     : ::|: :||:|: ::||:|:: |||  ||| ||::|||:|| |||||
Qy    148 TATQDA-DSSRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGM 206

Db    121 e 121
          |
Qy    207 E 207


RESULT    11
ID    ERR2_HUMAN      STANDARD;       PRT;   433 AA.
AC    P11475;
DT    01-OCT-1989 (REL. 12, CREATED)
DT    01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)
DT    01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE    STEROID HORMONE RECEPTOR ERR2.
GN    ERR2.
OS    HOMO SAPIENS (HUMAN).
OC    EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC    EUTHERIA; PRIMATES.
RN    [1]
RP    SEQUENCE FROM N.A.
RC    TISSUE=HEART;
RX    MEDLINE; 88122546.
RA    GIGUERE V., YANG N., SEGUI P., EVANS R.M.;
RL    NATURE 331:91-94(1988).
CC    -!- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC    -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC        NUCLEAR HORMONE RECEPTORS.
DR    EMBL; X51417; G36611; -.
DR    PIR; B29345; B29345.
DR    HSSP; P06536; 1GDC.
DR    PROSITE; PS00031; STEROID_FINGER.
KW    RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW    ZINC-FINGER.
FT    DNA_BIND    103    168        C4-TYPE ZINC FINGERS (TWO).
FT    ZN_FING     103    123        C4-TYPE.
FT    ZN_FING     139    163        C4-TYPE.
SQ    SEQUENCE   433 AA;  48287 MW;  673BF88F CRC32;

  Query Match             28.2%;  Score 486;  DB 3;  Length 433;
  Best Local Similarity   36.1%;  Pred. No. 9.73e-72;
  Matches    83;  Conservative   63;  Mismatches  70;  Indels  14;  Gaps   8;

Db    216 llvaepdklyamppddvpeggdikalttlcdladrelvfliswakhipgfsnltlgdqmsl 275
          || |||   :       |     |   :       : :|  |||:||| :||||| |||| :|:| ||: |
Qy      5 LLEAEPPHVLISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFDQVRL 63

Db    276 lqsawmeillilqivyrslpyddklayaedyimdeehsrlv-glleiyrailqlvrrykkl 334
          |:| |||:|::|:: ||:     || :| | ::| : :: | |:||:: :|     |:: |
Qy     64 LESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSRFREL 123

Db    335 kvekeefvmlkalalansdsmy--i---enleavqklqdilh---ealq-dyels--qrh 383
          |::  |:: :||: | || |||  :     :: :. .||  ||.   .||       |    .
Qy    124 KLQHKEYLCVKAMILLNS-SMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKSGISSQ 182
```

```
Db    384 eeprragklltlpllrqtaakavqhfysvklqgkvpmhklflemleakv 433
          ... | ..||. |. .|... |...|. ..|     ||.  |.||||.| |
Qy    183 QQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNAHV 232


RESULT    12
ID    ERR1_HUMAN      STANDARD;       PRT;   521 AA.
AC    P11474;
DT    01-OCT-1989 (REL. 12, CREATED)
DT    01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)
DT    01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE    STEROID HORMONE RECEPTOR ERR1.
GN    ERR1.
OS    HOMO SAPIENS (HUMAN).
OC    EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC    EUTHERIA; PRIMATES.
RN    [1]
RP    SEQUENCE FROM N.A.
RC    TISSUE=KIDNEY;
RX    MEDLINE; 88122546.
RA    GIGUERE V., YANG N., SEGUI P., EVANS R.M.;
RL    NATURE 331:91-94(1988).
CC    -!- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC    -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC        NUCLEAR HORMONE RECEPTORS.
DR    EMBL; X51416; G36609; -.
DR    PIR; A29345; A29345.
DR    HSSP; P06536; 1GDC.
DR    PROSITE; PS00031; STEROID_FINGER.
KW    RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW    ZINC-FINGER.
FT    DNA_BIND    176    241        C4-TYPE ZINC FINGERS (TWO).
FT    ZN_FING     176    196        C4-TYPE.
FT    ZN_FING     212    236        C4-TYPE.
SQ    SEQUENCE   521 AA;  55742 MW;  E2D723EA CRC32;

  Query Match             25.7%;  Score 444;  DB 3;  Length 521;
  Best Local Similarity   36.4%;  Pred. No. 3.13e-63;
  Matches    84;  Conservative   51;  Mismatches  89;  Indels   7;  Gaps   4;

Db    294 lvshllvvepeklyampdpagpdghlpavatlcdlfdreivvtiswaksipgfsslslsd 353
          ||  ||  ||  :   :: |::|        : . :|  | |:|:|  ||||| ||||  ||| |
Qy      1 LVLTLLEAEPPHVL-ISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 59

Db    354 qmsvlqsvwmevlvlgvaqrslplqdelafaedlvld-eeqaraaglgelgaallqlvrr 412
          |: :|:| ||||||::|:  ||:       | || ||||| :||      |: |:    ||      |
Qy     60 QVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSR 119

Db    413 lqalrlereeyvllkalalansdsvhiedeprlwsscekllh--eal---leyeaqraqp 467
          :: |:|:: ||: :||: | ||       :       :     |   || |   :|:    |       :::|
Qy    120 FRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKSGI 179

Db    468 qqgaerrragrllltlpllrqtagkvlahfyqvklegkvpmhklflemlea 518
          ::   :   | : ||: |: :|::: | : |:   :|      ||:   |:||||:|
Qy    180 SSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNA 230


RESULT    13
ID    PRGR_HUMAN      STANDARD;       PRT;   933 AA.
AC    P06401;
DT    01-JAN-1988 (REL. 06, CREATED)
DT    01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT    01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE    PROGESTERONE RECEPTOR (PR).
GN    PGR.
OS    HOMO SAPIENS (HUMAN).
OC    EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC    EUTHERIA; PRIMATES.
RN    [1]
RP    SEQUENCE FROM N.A.
RX    MEDLINE; 90228361.
RA    KASTNER P., KRUST A., TURCOTTE B., STROPP U., TORA L., GRONEMEYER H.,
RA    CHAMBON P.;
```

```
RL   EMBO J. 9:1603-1614(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 87184565
RA   MISRAHI M., ATGER M., D'AURIOL L., LOOSFELT H., MERIEL C.,
RA   FRIDLANSKY F., GUIOCHON-MANTEL A., GALIBERT F., MILGROM E.;
RL   BIOCHEM. BIOPHYS. RES. COMMUN. 143:740-748(1987).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- ALTERNATIVE PRODUCTS: THERE ARE TWO FORMS OF THE PROGESTERONE
CC       RECEPTOR: FORM A (THAT STARTS WITH MET-1) AND FORM B (THAT STARTS
CC       WITH MET-165).
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; X51730; G35652; -.
DR   EMBL; M15716; G189935; -.
DR   PIR; A03245; QRHUP.
DR   PIR; S09971; S09971.
DR   HSSP; P06536; 1GDC.
DR   MIM; 264080; -.
DR   PROSITE; PS00031; STEROID_FINGER.
DR   TRANSFAC; T00696; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING; PHOSPHORYLATION; ALTERNATIVE SPLICING.
FT   DOMAIN        1    566       MODULATING, PRO-RICH.
FT   DNA_BIND    567    632       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     567    587       C4-TYPE.
FT   ZN_FING     603    627       C4-TYPE.
FT   DOMAIN      681    933       STEROID-BINDING.
FT   DOMAIN      183    187       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT   MOD_RES      41     41       PHOSPHORYLATION (BY CK).
FT   MOD_RES     227    227       PHOSPHORYLATION (BY CK).
FT   MOD_RES     232    232       PHOSPHORYLATION (BY CK).
FT   MOD_RES     552    552       PHOSPHORYLATION (BY CK).
FT   MOD_RES     793    793       PHOSPHORYLATION (BY CK).
FT   VARSPLIC      1    164       MISSING (IN FORM B).
FT   CONFLICT    119    119       F -> G (IN REF. 2).
FT   CONFLICT    159    159       I -> V (IN REF. 2).
FT   CONFLICT    344    344       T -> S (IN REF. 2).
FT   CONFLICT    660    660       L -> V (IN REF. 2).
SQ   SEQUENCE   933 AA;  99042 MW;  CD1124EB CRC32;

  Query Match            19.6%;  Score 338;  DB 7;  Length 933;
  Best Local Similarity  28.8%;  Pred. No. 3.30e-42;
  Matches   59;  Conservative   61;  Mismatches   75;  Indels   17;  Gaps   17;

Db    687 linllmsiepdvivaqhdrtkpdtssslltslnqlgerqllsvvkwskslpgfrnlhidd 746
          |:  |:  ||  :  :: :: | | :|:: ||  |::::|: :: |:| :||| :| : |
Qy      1 LVLTLLEAEPPHVLISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 59

Db    747 qitliqyswmslmvfglgwrsykhvsgqmlyfapdlilneqrmk-essfyslcltmwqip 805
          |: |::  || ::: || ||:  ' :| :: |||||:|: :  |    ::  :   :
Qy     60 QVPLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDMLLATT 117

Db    806 qefvklqvsqeeflcmkvllllntiplegllsqtq-feemis-syl-relikalg-lrqk 861
             |  |  :  : |:||:| ::'||: ::  | : ||   :  |   : :    : |: :  |
Qy    118 SPRPFLKLQHKFYICVKAMILLNS-SMYPLVTATQDADSSPKLAHLLNAVTDALVWVIAK 176

Db    862 qvvsssqrfyqltklldnlhdlvkq 886
          : :|| |:  :|::||  |    |::
Qy    177 SGISSQQQSMRLANLL-MLLSHVRH 200


RESULT   14
ID   PRGR_MOUSE     STANDARD;      PRT;   923 AA.
AC   Q00175;
DT   01-DEC-1992 (REL. 24, CREATED)
DT   01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   PROGESTERONE RECEPTOR (PR).
GN   PGR OR PR.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 91299759.
RA   SCHOTT D.R., SHYAMALA G., SCHNEIDER W., PARRY G.;
RL   BIOCHEMISTRY 30:7014-7020(1991).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; M68915; G200472; -.
DR   PIR; A39596; A39596.
DR   HSSP; P06536; 1GDC.
DR   PROSITE; PS00031; STEROID_FINGER.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
FT   DOMAIN        1    556       MODULATING, PRO-RICH.
FT   DNA_BIND    557    622       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     557    577       C4-TYPE.
FT   ZN_FING     593    617       C4-TYPE.
FT   DOMAIN      671    923       STEROID-BINDING.
FT   DOMAIN      184    188       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
SQ   SEQUENCE   923 AA;  99073 MW;  EEDFF41A CRC32;

  Query Match            19.5%;  Score 336;  DB 7;  Length 923;
  Best Local Similarity  27.5%;  Pred. No. 8.08e-42;
  Matches   57;  Conservative   66;  Mismatches   70;  Indels   14;  Gaps   14;

Db    677 lillmslepdvlyaghdntkpdtssslltslnqlgerqllsvvkwskslpgfrnlhidd 736
          |.  |.  ||   .   .. .. | | .|.. ||  |::::|: :: |:| :||| :: : |
Qy      1 LVLTLLEAEPPHVLISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 59

Db    737 qitliqyswmslmvfglgwrsykhvsgqmlyfapdlilne-q-rmke-lsfyslcltmw 792
          |. |..  || ::: ||' |||  | .| .. ||||||:|. .. .  :   | :: : | :
Qy     60 QVPLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDM-L-L- 114

Db    793 qipqefvklqvtheeflcmkvllllnt-iplegl-sqsqfeemrssyirelikaiq-lr 849
               |  | : ! |:||:' ::||: : ||       :::   : :  :  :  |:  :
Qy    115 ATTSPRPFLKLQHKFYICVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVI 174

Db    850 qkgvvptsqrfyqltklldslhdlvkq 875
           |: ::: |:  :|::||  |    |::
Qy    175 AKSGISSQQQSMRLANLL-MLLSHVRH 200


RESULT   15
ID   PRGR_RABIT     STANDARD;      PRT;   930 AA.
AC   P06186;
DT   01-JAN-1988 (REL. 06, CREATED)
DT   01-JAN-1988 (REL. 06, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   PROGESTERONE RECEPTOR (PR).
GN   PGR.
OS   ORYCTOLAGUS CUNICULUS (RABBIT).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; LAGOMORPHA.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 87067449.
RA   LOOSFELT H., ATGER M., MISRAHI M., GUIOCHON-MANTEL A., MERIEL C.,
RA   LOGEAT F., BENAROUS R., MILGROM E.;
RL   PROC. NATL. ACAD. SCI. U.S.A. 83:9045-9049(1986).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; M14547; G165632; -.
DR   PIR; A25923; A25923.
DR   HSSP; P06536; 1GDC.
DR   PROSITE; PS00031; STEROID_FINGER.
DR   TRANSFAC; T00697; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
FT   DOMAIN          1    565       MODULATING, PRO-RICH.
FT   DNA_BIND      568    633       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING       568    588       C4-TYPE.
FT   ZN_FING       604    628       C4-TYPE.
FT   DOMAIN        678    930       STEROID-BINDING.
FT   DOMAIN        184    188       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
SQ   SEQUENCE   930 AA;  98666 MW;  EE133E86 CRC32;

  Query Match             19.5%;  Score 336;  DB 7;  Length 930;
  Best Local Similarity   27.9%;  Pred. No. 8.08e-42;
  Matches    57;  Conservative   62;  Mismatches  77;  Indels    8;  Gaps    8;

Db    684 linllmsiepdviyaqhdntkpdtssslltslnqlgerqllsvvkwskslpgfrnlhidd 743
          |:  |:  ||  :  :: :: | | :|:: ||  |::::|: :: |:| :||| :| : |
Qy      1 LVLTLLEAEPPHVLISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 59

Db    744 qitliqyswmslmvfqlqwreykhvsqqmlyfapdlilneqrmk-essfyslqltmwqip 802
          |: |::  || ::: || |||  | :| :: |||||:|: :  |    ::  :   :
Qy     60 QVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDMLLATT 117

Db    803 qefvklqvsqeeflcmkvlillnt i pleqlrsqsqfeemrssyirellkaig lrqkg 859
            |  | : : |:||:| ::|||: : ||     :::   : :  :  : |:  : |:
Qy    118 SRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKS 177

Db    860 vvsssqrfyqltklldnlhdlvkq 883
           :|| |:  :|::||  |   |::
Qy    178 GISSQQQSMRLANLL-MLLSHVRH 200

Search completed: Thu Dec 18 08:36:11 1997
Job time : 42 secs.
```

```
*****************************************************************************

                          MPSRCH                                    (TM)

*****************************************************************************

           Release 2 1D John F. Collins, Biocomputing Research Unit.
           Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
                   Distribution rights by IntelliGenetics, Inc

MPsrch_pp   protein - protein database search, using Smith-Waterman algorithm

Run on:          Thu Dec 18 08:34:01 1997;  MasPar time 9.77 Seconds
                                            689 205 Million cell updates/sec
Tabular output not generated

Title:           >US-08-826-361A-4
Description:     (1-233) from US08826361A.pep
Perfect Score:   1726
Sequence:        1 LVLTLLEAEPPHVLISRPSA.........CKNVVPVYDLLLEMLNAHVL 233

Scoring table:   PAM 150
                 Gap 11

Searched:        91006 seqs, 28888923 residues

Post-processing. Minimum Match 0%
                 Listing first 45 summaries

Database:        pir51
                   1:ann1 2:ann2 3:ann3 4:ann4 5:unann1 6:unann2 7:unann3
                   8:unann4 9:unann5 10:unann6 11:unann7 12:unann8
                   13:unann9 14:unann10 15:unenc 16:unrev

Statistics:      Mean 45.326;  Variance 111.522,  scale = 406

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                               SUMMARIES
             %
Result       Query
  No.  Score Match Length DB  ID              Description             Pred. No.
-------------------------------------------------------------------------------
    1   1142  66.2   589  3  QRCHE          estrogen receptor -   6.02e-166
    2   1122  65.0   599  3  QRMSE          estrogen receptor -   1.50e-162
    3   1118  64.8   600  3  QRPTE          estrogen receptor -   7.17e-162
    4   1116  64.7   595  3  QRHUE          estrogen receptor -   1.57e-161
    5   1106  64.1   586  3  QRXLE          estrogen receptor -   7.80e-160
    6   1066  61.8   535  7  S58224         oestrogen receptor -  4.72e-153
    7   1048  60.7   574  7  A37197         estrogen receptor -   5.29e-150
    8    920  53.3   701 16  S64737         WOK estrogen recepto  2.16e-128
    9    579  33.5   133  7  S26595         estrogen receptor -   1.08e-71
   10    499  28.9   121 13  I67419         estradiol receptor -  1.08e-58
   11    486  28.2   432 13  P29345         steroid hormone rece  1.35e-56
   12    470  27.2   433 14  S59087         estrogen receptor re  5.05e-54
   13    444  25.7   521 13  A29345         steroid hormone rece  7.32e-50
   14    340  19.7   923 14  I53280         progesterone recepto  1.65e-33
   15    338  19.6   933  3  QRHUP          progesterone recepto  3.35e-33
   16    336  19.5   923  7  A39596         progesterone recepto  6.82e-33
   17    336  19.5   930  7  A25923         progesterone recepti  6.82e-33
   18    322  18.7   758 16  S60586         glucocorticoid recep  9.67e-31
   19    314  18.2   776  3  S44047         glucocorticoid recep  1.62e-29
   20    311  18.0   783  3  A25691         glucocorticoid recep  4.64e-29
   21    307  17.8   742  3  QRHUGB         glucocorticoid recep  1.89e-28
   22    307  17.8   777  3  QRHUGA         glucocorticoid recep  1.89e-28
   23    304  17.6   771  7  A54273         glucocorticoid recep  5.39e-28
   24    299  17.3   533 13  S37781         retinoid X receptor   3.09e-27
   25    296  17.1   410 14  [illegible]    retinoic acid recept  9.78e-27
   26    296  17.1   446 14  A34418         H 2 region II bindin  9.78e-27
   27    296  17.1   448 14  D41727         retinoid X receptor   9.78e-27
   28    296  17.1   451 14  A41651         retinoic acid recept  9.78e-27
   29    293  17.0   795  3  QRRTG          glucocorticoid recep  2.49e-26
   30    286  16.6   379 16  I50514         retinoid X receptor   2.83e-25
   31    286  16.6   462 13  S09592         retinoic acid recept  2.83e-25
   32    286  16.6   466 14  C41727         retinoid X receptor   2.83e-25
   33    286  16.6   467 14  A47278         retinoid X receptor   2.83e-25
   34    286  16.6   467 14  S26668         retinoic acid recept  2.83e-25
   35    286  16.6   467 13  A43781         retinoid-X-receptor-  2.83e-25
   36    285  16.5   786  7  A35466         progesterone recepto  2.90e-25
   37    285  16.5   984 13  A29513         mineralocorticoid re  2.90e-25
   38    284  16.5   441 16  I50515         retinoid X receptor   5.65e-25
   39    284  16.5   981 14  A41401         mineralocorticoid re  5.65e-25
   40    283  16.4   340 14  B40702         retinoid-X-receptor-  7.98e-25
   41    283  16.4   463 14  B41727         retinoid-X receptor-  7.98e-25
   42    282  16.3   385 14  I52158         orphan nuclear recep  1.13e-24
   43    279  16.2   385 13  S48074         nuclear receptor Tlx  3.17e-24
   44    280  16.2   462 14  S26670         retinoic acid recept  2.25e-24
   45    279  16.2   488 13  C41377         retinoid receptor al  3.17e-24


                               ALIGNMENTS

RESULT     1
ENTRY               QRCHE      #type complete
TITLE               estrogen receptor - chicken
ORGANISM            #formal_name Gallus gallus #common_name chicken
DATE                06-Mar-1992 #sequence_revision 14-Jul-1994 #text_change
                      15-Jun-1996
ACCESSIONS          A40914; S07192
REFERENCE           A40914
   #authors         Maxwell, B.L.; McDonnell, D.P.; Conneely, O.M.; Schulz, T.Z.;
                      Greene, G.L.; O'Malley, B.W.
   #journal         Mol. Endocrinol. (1987) 1:25-35
   #title           Structural organization and regulation of the chicken
                      estrogen receptor.
   #cross-references MUID.88318621
   #accession       A40914
      ##status           not compared with conceptual translation
      ##molecule_type mRNA
      ##residues         1-589 ##label MAX
REFERENCE           S07192
   #authors         Krust, A.; Green, S.; Argos, P.; Kumar, V.; Walter, P.;
                      Bornert, J.M.; Chambon, P.
   #journal         EMBO J (1986) 5:891-897
   #title           The chicken oestrogen receptor sequence: homology with v-erbA
                      and the human oestrogen and glucocorticoid receptors.
   #cross-references MUID.86247578
   #accession       S07192
      ##status           preliminary
      ##molecule_type mRNA
      ##residues         1-256,'E',258-589 ##label KRU
      ##cross-references EMBL.X03805
COMMENT    The steroid hormones and their receptors are involved in the
             regulation of eukaryotic gene expression and affect cellular
             proliferation and differentiation in target tissues.
COMMENT    In the absence of ligand, steroid hormone receptors are thought to
             be weakly associated with nuclear components; hormone binding
             greatly increases receptor affinity. The hormone-receptor complex
             appears to recognize discrete DNA sequences upstream of
             transcriptional start sites.
CLASSIFICATION    #superfamily estrogen receptor, erbA transforming protein
                      homology
KEYWORDS            DNA binding; nucleus; phosphoprotein; steroid binding;
                      steroid hormone receptor; transcription regulation; zinc
                      finger
FEATURE
   1-173                #domain amino terminal #label NH2\
     174-265
     177-450           #domain DNA binding #status predicted #label DNA\
     177-200           #domain erbA transforming protein homology #label ERBA\
     213-235           #region zinc finger CCCC motif\
     250-265           #region zinc finger CCCC motif\
     294-546           #region nuclear location signal\
     179,182,196,199   #domain steroid binding #status predicted #label STB\
     215,221,231,234   #binding_site zinc (Cys) #status predicted\
     230,299           #binding_site zinc (Cys) #status predicted\
                       #binding_site phosphate (Ser) (covalent) #status
SUMMARY                  predicted
                   #length 589  #molecular-weight 66773  #checksum 793

  Query Match             66.2%;  Score 1142;  DB 3;  Length 589;
  Best Local Similarity   62.1%;  Pred. No. 6.02e-166;
  Matches  146;  Conservative  53;  Mismatches  34;  Indels  2;  Gaps  2;

Db     309 mvsalleaeppivyseydpnrpfneasmmtlltnladrelvhminwakrvpgfvdltlhd 368
            :| :||||||| |          |: || |||||  ||:|||:||||||:|||::||||.|:| |
Qy       1 LVLTLLEAEPPHVLIS-RPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 59

Db     369 qvhllecawleilmiglvwrsmehpgkllfapnllldrnqgkcvegmveifdmllataar 428
           ||:|||  |:|:||:||:|||::|||||:|||:|:|||::||||||::|||||||||:.|
Qy      60 QVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSR 119

Db     429 frmmnlqgeefvclksiillnsgvyttlsstlksleerdyihrvldkitdtlihlmaksg 488
           || ::||  |::|:|::||||||::|  ....:|   .   |   | :|: :||:|: ::||||
Qy     120 FRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAH-LLNAVTDALVWVIAKSG 178

Db     489 lslqqhrrlaqlllilshirhmsnkgmehlynmkcknvvplydlllemldahrl 543
           :| |||  ||| ||::|||:|| |||||||| ||||||||||:|||||||||:|| |
Qy     179 ISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNAHVL 233

RESULT      2
ENTRY
TITLE                 QRMSE       #type complete
ORGANISM              estrogen receptor - mouse
DATE                  #formal_name Mus musculus #common_name house mouse
                      06-Mar-1992 #sequence_revision 14-Jul-1994 #text_change
                        15-Jun-1996
ACCESSIONS            A40061
REFERENCE             A40061
   #authors           White, R.; Lees, J.A.; Needham, M.; Ham, J.; Parker, M.
   #journal           Mol. Endocrinol. (1987) 1:735-744
   #title             Structural organization and expression of the mouse estrogen
                        receptor.
   #cross-references MUID:91042558
   #accession        A40061
      ##molecule_type mRNA
      ##residues        1-599 ##label WHI
      ##cross-references GB:M38651
COMMENT      The steroid hormones and their receptors are involved in the
               regulation of eukaryotic gene expression and affect cellular
               proliferation and differentiation in target tissues.
COMMENT      In the absence of ligand, steroid hormone receptors are thought to
               be weakly associated with nuclear components; hormone binding
               greatly increases receptor affinity. The hormone-receptor complex
               appears to recognize discrete DNA sequences upstream of
               transcriptional start sites.
CLASSIFICATION   #superfamily estrogen receptor; erbA transforming protein
                   homology
KEYWORDS         DNA binding; nucleus; phosphoprotein; steroid binding;
                   steroid hormone receptor; transcription regulation; zinc
                   finger
FEATURE
   1-183
   184-275               #domain amino-terminal #label NH2\
   187-460               #domain DNA binding #status predicted #label DNA\
   187-210               #domain erbA transforming protein homology #label ERBA\
   223-245               #region zinc finger CCCC motif\
   260-275               #region zinc finger CCCC motif\
   304-556               #region nuclear location signal\
   189,192,206,209       #domain steroid binding #status predicted #label STB\
   225,231,241,244       #binding_site zinc (Cys) #status predicted\
                         #binding_site zinc (Cys) #status predicted\
```

```
     240,309
                         #binding_site phosphate (Ser) (covalent) #status
SUMMARY                    predicted
                   #length 599  #molecular-weight 66954  #checksum 4749

  Query Match             65.0%;  Score 1122;  DB 3;  Length 599;
  Best Local Similarity   60.0%;  Pred. No. 1.50e-162;
  Matches  141;  Conservative  59;  Mismatches  33;  Indels  2;  Gaps  2;

Db     319 mvsalldaeppmiyseydpsrpfseasmmglltnladrelvhminwakrvpgfgdlnlhd 378
            :| :||:|||| :         || ||:|||||  ||:|||:||||||:|||::|||  :|:| |
Qy       1 LVLTLLEAEPPHVLIS-RPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 59

Db     379 qvhilecawleilmiglvwrsmehpgkllfapnllldrnqgkcveqmveifdmllatssr 438
           ||:|||  |:|:||.||:|||::|||||:|||:|:|||·:||||||::||||||||||:||
Qy      60 QVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSR 119

Db     439 frmmnlqgeefvclksiillnsgvytflsstlksleekdhihrvldkitdtlihlmakaq 498
           || ::||  |::|:|..||||·:| ::..|  : :      : ::|: :||:|: ::||:|
Qy     120 FRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSRKLAHLLNAVTDALVWVIAKSG 178

Db     499 ltlqqqhrrlaqlllilshirhmsnkgmehlynmkcknvvplydlllemldahrl 553
           :: |||  ||| ||..|||:|| |||||||| ||||||||||:|||||||||:|| |
Qy     179 ISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNAHVL 233

RESULT      3
ENTRY
TITLE                 QRRTE       #type complete
ORGANISM              estrogen receptor - rat
DATE                  #formal_name Rattus norvegicus #common_name Norway rat
                      30-Sep-1991 #sequence_revision 14-Jul-1994 #text_change
                        15-Jun-1996
ACCESSIONS            S07379; S16731
REFERENCE             S07379
   #authors           Koike, S.; Sakai, M.; Muramatsu, M.
   #journal           Nucleic Acids Res. (1987) 15:2499-2513
   #title             Molecular cloning and characterization of rat estrogen
                        receptor cDNA.
   #cross-references MUID:87174780
   #accession        S07379
      ##molecule_type mRNA
      ##residues        1-600 ##label KOI
      ##cross-references EMBL:Y00102
REFERENCE             S16731
   #authors           Maggi, A.M.A.
   #submission        submitted to the EMBL Data Library, June 1991
   #accession         S16731
      ##molecule_type mRNA
      ##residues        1-487,'T',489-600 ##label MAG
      ##cross-references EMBL:X61098
COMMENT      The steroid hormones and their receptors are involved in the
               regulation of eukaryotic gene expression and affect cellular
               proliferation and differentiation in target tissues.
COMMENT      In the absence of ligand, steroid hormone receptors are thought to
               be weakly associated with nuclear components; hormone binding
               greatly increases receptor affinity. The hormone-receptor complex
               appears to recognize discrete DNA sequences upstream of
               transcriptional start sites.
CLASSIFICATION   #superfamily estrogen receptor; erbA transforming protein
                   homology
KEYWORDS         DNA binding; nucleus; phosphoprotein; steroid binding;
                   steroid hormone receptor; transcription regulation; zinc
                   finger
FEATURE
   1-184
   185-276               #domain amino-terminal #label NH2\
   188-461               #domain DNA binding #status predicted #label DNA\
   188-211               #domain erbA transforming protein homology #label ERBA\
   224-246               #region zinc finger CCCC motif\
   261-276               #region zinc finger CCCC motif\
   305-557               #region nuclear location signal\
   190,193,207,210       #domain steroid binding #status predicted #label STB\
   226,232,242,245       #binding_site zinc (Cys) #status predicted\
   241,310               #binding_site zinc (Cys) #status predicted\
                         #binding_site phosphate (Ser) (covalent) #status
```

```
                           predicted
SUMMARY          #length 600  #molecular-weight 67029  #checksum 4599

  Query Match              64.8%;  Score 1118;  DB 3;  Length 600;
  Best Local Similarity   59.6%;  Pred. No. 7.17e-162;
  Matches   140;  Conservative   59;  Mismatches  34;  Indels   2;  Gaps   2;

Db     320 mvsalldaeppliyseydpsrpfseasmmglltnladrelvhminwakrvpgfgdlnlhd 379
           :| ::|:||||| :     || ||.||||| ||.||'.||||||.|||:.|||' :|:| |
Qy       1 LVLTLLEAEPPHVLIS-RPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 59

Db     380 qvhllecawleilmiglvwrsmehpgkllfapnllldrnqgkcvegmveifdmllatssr 439
           ||:||| |:|:||:||:|||:|||||:|||:|:||::||||||::||||||||||:||
Qy      60 QVRLLFSCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSR 119

Db     440 frmmnlqgeefvclksiillnsgvytflsstlksleekdhihrvldkitdtlihlmakag 499
           || ::||  |::|:|::'||||::| ...:|  .  .     . ..|. . |.|. .:|||:|
Qy     120 FRFLKIQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSRKLAHLLNAVTDALVWVIAKSG 178

Db     500 ltlqqqhrrlaqllilshirhmsnkgmehlynmkcknvvplydlllemldahrl 554
           :: |||  ||| |:::||:|| |||||||| |||:||||:|||||||||:|| |
Qy     179 ISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNAHVL 233


RESULT     4
ENTRY            QRHUE       #type complete
TITLE            estrogen receptor - human
ALTERNATE_NAMES  hER
ORGANISM         #formal_name Homo sapiens #common_name man
DATE             28-May-1986 #sequence_revision 28-May-1986 #text_change
                   15-Nov-1996
ACCESSIONS       A94284; A93376; A43021; S27143; S34000; A41925; B41925;
                   A03244; C41925; D41925
REFERENCE        A94284
   #authors      Greene, G.L.; Gilna, P.; Waterfield, M.; Baker, A.; Hort, Y.;
                   Shine, J.
   #journal      Science (1986) 231:1150-1154
   #title        Sequence and expression of human estrogen receptor
                   complementary DNA.
   #cross-references MUID:86122927
   #accession    A94284
      ##molecule_type mRNA
      ##residues      1-595 ##label GR1
      ##cross-references GB:M12674
REFERENCE        A93376
   #authors      Green, S.; Walter, P.; Kumar, V.; Krust, A.; Bornert, J.M.;
                   Argos, P.; Chambon, P.
   #journal      Nature (1986) 320:134-139
   #title        Human oestrogen receptor cDNA: sequence, expression and
                   homology to v-erb-A.
   #cross-references MUID:86146892
   #accession    A93376
      ##molecule_type mRNA
      ##residues      1-595 ##label GR2
      ##cross-references GB:X03635
REFERENCE        A43021
   #authors      Ponglikitmongkol, M.; Green, S.; Chambon, P.
   #journal      EMBO J. (1988) 7:3385-3388
   #title        Genomic organization of the human oestrogen receptor gene
   #accession    A43021
      ##molecule_type DNA
      ##residues      143-161,205-225,244-264,356-374,'S',376,402-412,447-460,
                        'P',462-467,508-528 ##label PON
REFERENCE        S27140
   #authors      Keaveney, M.; Klug, J.; Gannon, F.
   #journal      DNA Seq. (1992) 2:347-358
   #title        Sequence analysis of the 5' flanking region of the human
                   estrogen receptor gene.
   #accession    S27143
      ##status        preliminary; translation not shown
      ##molecule_type DNA
      ##residues      1-115 ##label KEA
      ##cross-references EMBL:X62462
REFERENCE        S34000
   #authors      Pfeffer, U.; Fecarotta, E.; Castagnetta, L.; Vidali, G.
   #journal      Cancer Res. (1993) 53:741-743
   #title        Estrogen receptor variant messenger RNA lacking exon 4 in
                   estrogen-responsive human breast cancer cell lines.
   #accession    S34000
      ##status        preliminary
      ##molecule_type mRNA
      ##residues      216-254,367-399,'G',401-434 ##label PFE
      ##cross-references EMBL:X73067
      ##note          the authors translated the codon GGG for residue 400 as
                        Val
REFERENCE        A41925
   #authors      Dotzlaw, H.; Alkhalaf, M.; Murphy, L.C.
   #journal      Mol. Endocrinol. (1992) 6:773-785
   #title        Characterization of estrogen receptor variant mRNAs from
                   human breast cancers.
   #cross-references MUID:92293154
   #accession    A41925
      ##molecule_type mRNA
      ##residues      1-214,'ELPTLC' ##label DOT
      ##cross-references NCBIN:106580
      ##experimental_source clone 4; breast cancer
      ##note          sequence has been revised after extraction from NCBI
                        backbone
      ##note          the complete sequence of neither the nucleotide nor the
                        protein is shown in this paper
   #accession    B41925
      ##status        significant sequence differences
      ##molecule_type mRNA
      ##cross-references NCBIN:106597; GB:M69296
      ##experimental_source clone 24; breast cancer
COMMENT    The steroid hormones and their receptors are involved in the
             regulation of eukaryotic gene expression and affect cellular
             proliferation and differentiation in target tissues.
COMMENT    In the absence of ligand, steroid hormone receptors are thought to
             be weakly associated with nuclear components; hormone binding
             greatly increases receptor affinity. The hormone-receptor complex
             appears to recognize discrete DNA sequences upstream of
             transcriptional start sites.
GENETICS
   #gene         GDB:ESR
      ##cross-references GDB:119120
   #map_position 6q25.1-6q25.1
   #introns      151/2; 215/1; 254/1; 366/1; 412/1; 457/1; 518/2
CLASSIFICATION   #superfamily estrogen receptor; erbA transforming protein
                   homology
KEYWORDS         DNA binding; nucleus; phosphoprotein; steroid hormone
                   receptor; transcription regulation; zinc finger
FEATURE
   1-120                  #domain amino-terminal #label NH2\
   121-299                #domain DNA binding #status predicted #label DNA\
   183-456                #domain erbA transforming protein homology #label ERBA\
   185-205                #region zinc finger CCCC motif\
   221-245                #region zinc finger CCCC motif\
   300-595                #domain steroid binding #status predicted #label STB\
   236,305                #binding_site phosphate (Ser) (covalent) #status
                            predicted
SUMMARY          #length 595  #molecular-weight 66258  #checksum 4317

  Query Match              64.7%;  Score 1116;  DB 3;  Length 595;
  Best Local Similarity   59.7%;  Pred. No. 1.57e-161;
  Matches   141;  Conservative   61;  Mismatches  30;  Indels   4;  Gaps   3;

Db     315 mvsalldaepp-ilyseydptrpfseasmmglltnladrelvhminwakrvpgfvdltlh 373
           :| :||:|||| :| |    |: ||:'|||| ||:|||:||||||:|||::||'|':|:!
Qy       1 LVLTLLEAEPPHVLISR--PSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLF 58

Db     374 dqvhllecawleilmiglvwrsmehpgkllfapnllldrnqgkcvegmveifdmllatss 433
           '||:||| |:|:||:||:|||:|| ||:|||:|:||||::||||||:::|'| |||':
Qy      59 DQVRLLFSCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTS 118

Db     434 rfrmmnlqgeefvclksiillnsgvytflsstlksleekdhihrvldkitdtlihlmaka 493
```

```
             ||| ::||  |::|:|::||||||::| ::::|   : :      : ::|: :||:|: ::||:
Qy     119 RFRELKIQHKFYLCVKAMILLNSSMYPLVTATQDA-DSSRKLAHLLNAVTDALVWVIAKS 177

Db     494 qltlqqhqrlaqllllilshirhmsnkgmehlysmkcknvvplydlllemldahrl 549
             |:: |||  ||| ||::|||:|| |||||||| :||||||||:|||||||||:|| |
Qy     178 GISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNAHVL 233


RESULT     5
ENTRY              QRXLE       #type complete
TITLE              estrogen receptor   African clawed frog
ORGANISM           #formal_name Xenopus laevis #common_name African clawed frog
DATE               29-Feb-1992 #sequence_revision 14-Jul-1994 #text_change
                     15-Jun-1996
ACCESSIONS         A40907
REFERENCE          A40907
   #authors        Weiler, I.J.; Lew, D.; Shapiro, D.J.
   #journal        Mol. Endocrinol. (1987) 1:355-362
   #title          The Xenopus laevis estrogen receptor: sequence homology with
                     human and avian receptors and identification of multiple
                     estrogen receptor messenger ribonucleic acids.
   #cross-references MUID:90331927
   #accession      A40907
      ##molecule_type mRNA
      ##residues       1-586 ##label WEI
      ##cross-references GB:L20735
COMMENT     The steroid hormones and their receptors are involved in the
              regulation of eukaryotic gene expression and affect cellular
              proliferation and differentiation in target tissues.
COMMENT     In the absence of ligand, steroid hormone receptors are thought to
              be weakly associated with nuclear components; hormone binding
              greatly increases receptor affinity. The hormone-receptor complex
              appears to recognize discrete DNA sequences upstream of
              transcriptional start sites.
CLASSIFICATION     #superfamily estrogen receptor; erbA transforming protein
                     homology
KEYWORDS           DNA binding; nucleus; phosphoprotein; steroid binding;
                     steroid hormone receptor; transcription regulation; zinc
                     finger
FEATURE
   1-174                  #domain amino-terminal #label NH2\
   175-266                #domain DNA binding #status predicted #label DNA\
   178-448                #domain erbA transforming protein homology #label ERBA\
   178-201                #region zinc finger CCCC motif\
   214-237                #region zinc finger CCCC motif\
   251-266                #region nuclear location signal\
   292-544                #domain steroid binding #status predicted #label STB\
   180,183,197,200        #binding_site zinc (Cys) #status predicted\
   216,222,232,235        #binding_site zinc (Cys) #status predicted\
   231                    #binding_site phosphate (Ser) (covalent) #status
                            predicted
SUMMARY            #length 586  #molecular-weight 66080  #checksum 9991

  Query Match             64.1%;  Score 1106;  DB 3;  Length 586;
  Best Local Similarity  58.8%;  Pred. No. 7.80e-160;
  Matches   137;  Conservative   61;  Mismatches  33;  Indels   2;  Gaps   2;

Db     307 llsalmeaeplvysehdstkplseasmmtlltnladrelvhminwakrvpgfvdltlhd 366
             |: :|:||||:| |     : |: |::|||||  ||:|||:||||||:|||::||||:|:| |
Qy       1 LVLTLLEAEPPHVLISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 59

Db     367 qvhllecawleilmvqliwrsvehpqklsfapnllldrnqgrcveglveifdmlvttatr 426
             ||:|||  |:|:||:||:|||:|||||  |||:|:|||::|:`|||::||||||||::|::|
Qy      60 QVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSR 119

Db     427 frmmnlraeeficlksiillnsgvytflsstleslddtdlihildkiidtlvhfmaksg 486
             || ::|:  |::|:|::|||||::| ::::|  ··     | |· |· · |·||  ·||||
Qy     120 FRELKLQHKEYLCVKAMILLNSSMYPLVIATQDADSSRKLAHL-LNAVTDALVWVIAKSG 178

Db     487 lslqqqrrlaqlllilshirhmsnkgmehlysmkcknvvplydlllemldah 539
             :| |||  ||| ||::|||:|| |||||||| :||||||||:|||||||||:||
Qy     179 ISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNAH 231
```

```
RESULT     6
ENTRY              S58224      #type fragment
TITLE              estrogen receptor - Atlantic salmon (fragment)
ORGANISM           #formal_name Salmo salar #common_name Atlantic salmon
DATE               13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change
                     26-Apr-1996
ACCESSIONS         S58224
REFERENCE          S58224
   #authors        Rogers, S.A.; Llewellyn, L.; Ramsurn, V.P.; Sweeney, G.E.;
                     Wigham, T.
   #submission     submitted to the EMBL Data Library, July 1995
   #accession      S58224
      ##status          preliminary
      ##molecule_type mRNA
      ##residues       1-535 ##label ROG
      ##cross-references EMBL:X89959
CLASSIFICATION     #superfamily erbA transforming protein homology
FEATURE
   183-382                #domain erbA transforming protein homology #label ERBA
SUMMARY            #length 535  #checksum 9963

  Query Match             61.8%;  Score 1066;  DB 7;  Length 535;
  Best Local Similarity  60.9%;  Pred. No. 4.72e-153;
  Matches   143;  Conservative   40;  Mismatches  49;  Indels   3;  Gaps   3;

Db     241 vlfllqqaeppalcsrgkvarpytevtmmtlltsmadkelvhmiawakkvpgfqelslhd 300
             || ||::  |  | ||    | |:|| :||  || :||||||||||:||||:|||  |||| |
Qy       2 VLILLEAEPPHVLISRPS-A-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 59

Db     301 qvqllesswlevlmigliwrsihcpgklifaqdlildrseqdcvegmaeifdmllatvsr 360
             ||:|||| |:||||:||:|||| |||||||| ||:||| || ||||: |||||||| ||
Qy      60 QVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSR 119

Db     361 frmlklkpeefvclkaiillnsqafsfcsnsveslhnssavesmldnitdalihhishsq 420
             || |||   |..|.||.|||||. ... : . :.  .|  .  :|: :|||: |: ||
Qy     120 FRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSRKLAHLLNAVTDALVWVIAKSG 178

Db     421 asvqqqprqvqlllllshirhmsnkgmehlysikcknkvplydlllemldqhrl 475
             | |||: |  ||:||||:|| |||||||| ::|||| ||:||||||||::| :
Qy     179 ISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNAHVL 233


RESULT     7
ENTRY              A37197      #type complete
TITLE              estrogen receptor - rainbow trout
ORGANISM           #formal_name Oncorhynchus mykiss #common_name rainbow trout
DATE               21-Jan-1992 #sequence_revision 31-Jan-1992 #text_change
                     03-Jun-1996
ACCESSIONS         A37197; A40070
REFERENCE          A37197
   #authors        Pakdel, F.; Le Gac, F.; Le Goff, P.; Valotaire, Y.
   #journal        Mol. Cell. Endocrinol. (1990) 71:195-204
   #title          Full-length sequence and in vitro expression of rainbow trout
                     estrogen receptor cDNA.
   #cross-references MUID:91006824
   #accession      A37197
      ##status          preliminary
      ##molecule_type mRNA
      ##residues       1-574 ##label PAK
      ##cross-references GB:M31559
REFERENCE          A40070
   #authors        Pakdel, F.; Le Guellec, C.; Vaillant, C.; Le Roux, M.G.;
                     Valotaire, Y.
   #journal        Mol. Endocrinol. (1989) 3:44-51
   #title          Identification and estrogen induction of two estrogen
                     receptors (ER) messenger ribonucleic acids in the rainbow
                     trout liver: sequence homology with other ERs.
   #cross-references MUID:89127284
   #accession      A40070
      ##status          preliminary
      ##molecule_type mRNA
      ##residues       150-574 ##label PA2
      ##cross-references GB:M31559
CLASSIFICATION    #superfamily estrogen receptor, erbA transforming protein
                    homology
KEYWORDS          DNA binding; nucleus, steroid hormone receptor, zinc finger
FEATURE
   145-421              #domain erbA transforming protein homology #label ERBA\
   147-167              #region zinc finger\
   183-207              #region zinc finger
SUMMARY           #length 574  #molecular-weight 62871  #checksum 1057

  Query Match             60.7%;  Score 1048;  DB 7;  Length 574;
  Best Local Similarity   60.4%;  Pred. No. 5.29e-150;
  Matches  142;  Conservative   42;  Mismatches  47;  Indels   4;  Gaps   4;

Db      281 vlfllqqtp-alcsrqkvarpytevtmmtlltsmadkelvhmiawakkvpgfqelslhd 339
            || ||::: |   | ||    | |:|| :||   || :||||||||:||||:||| |||| |
Qy        2 VLTLLEAEPPHVLISRPS-A-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 59

Db      340 qvqllesswlevlmiglwrsihcpgklifaqdlildrsegdcvegmaeifdmllatvsr 399
            |::|||| |:||||:||:||||  |||||| ||:||| || ||||: ||||||||| ||
Qy       60 QVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSR 119

Db      400 fqmlklkpeefvclkaiillnpgafsfcsnsveslhnssavesmldnitdalihhishsg 459
            |  |||   |::|:||:||||:: ::: : . ..   .:  .   .:. .||||.  |. ||
Qy      120 FRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSRKLAHLLNAVTDALVWVIAKSG 178

Db      460 asvqqqprrqaqlllllshirhmsnkgmehlysikcknkvplydlllemldahrl 514
             | ||: | : ||:||||:|| ||||||||| ::||||| ||:||||||||::| |
Qy      179 ISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNAHVL 233


RESULT    8
ENTRY             S64737     #type complete
TITLE             80K estrogen receptor - human
ORGANISM          #formal_name Homo sapiens #common_name man
DATE              28-Oct-1996 #sequence_revision 28-Oct-1996 #text_change
                    28-Oct-1996
ACCESSIONS        S64737
REFERENCE         S64737
   #authors       Pink, J.J.; Wu, S.Q.; Wolf, D.M.; Pilimoria, M.M.; Jordan,
                    V.C.
   #journal       Nucleic Acids Res. (1996) 24:962-969
   #title         A novel 80 kDa human estrogen receptor containing a
                    duplication of exons 6 and 7.
   #accession     S64737
      ##status          preliminary
      ##residues        1-701 ##label PIN
      ##cross-references EMBL:U47678
SUMMARY           #length 701  #molecular-weight 78430  #checksum 4951

  Query Match             53.3%;  Score 920;  DB 16;  Length 701;
  Best Local Similarity   55.5%;  Pred. No. 2.16e-128;
  Matches  116;  Conservative   59;  Mismatches  30;  Indels   4;  Gaps   3;

Db      315 mvsalldaepp-ilyseydptrpfseasmmglltnladrelvhminwakrvpgfvdltlh 372
            :| :||:|||| :| |    |: ||:|||||   ||:||:||||||||:|||::||||:|:|
Qy        1 LVLTLLEAEPPHVLISR--PSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLF 58

Db      374 dqvhllecawleilmiglvwrsmehpvkllfapnllldrnqgkcvegmveifdmllatss 433
            |||:|||  |:|:||:||:||||::|| ||:|||:|:||:||::||||||||::||||||||:|
Qy       59 DQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTS 118

Db      434 rfrmmnlqgeefvclksiillnsgvytflsstlksleekdhihrvldkitdtlihlmaka 493
            ||| ::||  |::|:|::||||||::| ::::|   : :     : ::|: :||:|: ::|::
Qy      119 RFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSRKLAHLLNAVTDALVWVIAKS 177

Db      494 gltlqqqhqrlaqlllilshirhmrnqgk 522
            |:: |||  ||| ||::|||:||  | |:
Qy      178 GISSQQQSMRLANLLMLLSHVRHASNKGM 206


RESULT    9
ENTRY             S26595     #type fragment
```

```
TITLE             estrogen receptor - bovine (fragment)
ORGANISM          #formal_name Bos primigenius taurus #common_name cattle
DATE              06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change
                    12-Apr-1995
ACCESSIONS        S26595
REFERENCE         S26595
   #authors       Hagen-Mann, K.; Mann, W.; Meyer, H.H.D.
   #submission    submitted to the EMBL Data Library, May 1992
   #accession     S26595
      ##status          preliminary
      ##molecule_type mRNA
      ##residues        1-133 ##label HAG
      ##cross-references EMBL:X66841
CLASSIFICATION    #superfamily estrogen receptor; erbA transforming protein
                    homology
SUMMARY           #length 133  #checksum 5431

  Query Match             33.5%;  Score 579;  DB 7;  Length 133;
  Best Local Similarity   52.7%;  Pred. No. 1.08e-71;
  Matches   69;  Conservative   43;  Mismatches  18;  Indels   1;  Gaps   1;

Db        1 ltlhdqvhllecawleilmiglvwrsmehpvkllfapnllldrnqgkcvegmveifdmll 60
            |.| |||:|||  |:|:||:||:||||::|| ||:|||:|:||:||::||||||||::||.::.
Qy       55 LSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLL 114

Db       61 atssrfrmmnlqgeefvclksiillnsgvytflsstlksleekdhihrvldkitdtlihl 120
            ||.|||| ..||  |::|.|..|||||..| ::::|   : :     : ::|: :||:: :
Qy      115 ATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSRKLAHLLNAVTDALVWV 173

Db      121 makagltlqqq 131
            :|::|:: |||
Qy      174 IAKSGISSQQQ 184


RESULT   10
ENTRY             I67419     #type fragment
TITLE             estradiol receptor - rhesus macaque (fragment)
ORGANISM          #formal_name Macaca mulatta #common_name rhesus macaque
DATE              04-Oct-1996 #sequence_revision 04-Oct-1996 #text_change
                    04-Oct-1996
ACCESSIONS        I67419
REFERENCE         I53287
   #authors       Chandrasekher, Y.A., Melner, M.H.; Nagalla, S.R.; Stouffer,
                    R.L.
   #journal       Endocrinology (1994) 135:307-314
   #title         Progesterone receptor, but not estradiol receptor, messenger
                    ribonucleic acid is expressed in luteinizing granulosa
                    cells and the corpus luteum in rhesus monkeys.
   #cross-references MUID:94293272
   #accession     I67419
      ##status          preliminary, translated from GB/EMBL/DDBJ
      ##molecule_type mRNA
      ##residues        1-121 ##label RES
      ##cross-references GB:S71040; NID:g547182; CDS_PID:g547183
SUMMARY           #length 121  #checksum 4189

  Query Match             28.9%;  Score 499;  DB 13;  Length 121;
  Best Local Similarity   52.1%;  Pred. No. 1.08e-58;
  Matches   63;  Conservative   38;  Mismatches  19;  Indels   1;  Gaps   1;

Db        1 lfapnllldrnqgkcvegmvesfdmllatssrfrmmnlqgeefvclksiillnsgvytfl 60
            :|||:|:|||. |||||  | ||||||||:|||| ::||  |::|:|::||||||::| ::
Qy       88 IFAPDLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLV 147

Db       61 sstlksleekdhihrvldkitdtlihlmakagltlqqqhrrlaqllllilshirhmsnkgm 120
            ::|   : :     : ::|: :||:|: ::|::|:: |||  ||| ||::|||:|| |||||
Qy      148 TATQDA-DSSRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGM 206

Db      121 e 121
            |
Qy      207 E 207
```

```
RESULT    11
ENTRY              B29345      #type complete
TITLE              steroid hormone receptor ERR2 precursor - human
ORGANISM           #formal_name Homo sapiens #common_name man
DATE               15-Dec-1988 #sequence_revision 15-Dec-1988 #text_change
                     03-Jun-1996
ACCESSIONS         B29345
REFERENCE          A29345
   #authors        Giguere, V.; Yang, N.; Segui, P.; Evans, R.M.
   #journal        Nature (1988) 331:91-94
   #title          Identification of a new class of steroid hormone receptors.
   #cross-references MUID:88122546
   #accession      B29345
      ##molecule_type mRNA
      ##residues        1-433 ##label GIG
      ##cross-references GB:X51417; GB:Y00290
CLASSIFICATION     #superfamily erbA transforming protein homology
KEYWORDS           DNA binding; zinc finger
FEATURE
   101-352              #domain erbA transforming protein homology #label ERBA\
   103-123              #region zinc finger\
   139-163              #region zinc finger
SUMMARY            #length 433  #molecular-weight 48287  #checksum 1133

  Query Match              28.2%;  Score 486;  DB 13;  Length 433;
  Best Local Similarity   36.1%;  Pred. No. 1.35e-56;
  Matches    83;  Conservative    63;  Mismatches   70;  Indels   14;  Gaps    8;

Db      216 llvaepdklyamppddvpeqdikalttlcdladrelvflliswakhipgfsnltlgdqmsl 275
            || |||   :      |    |   :     : :|  |||:||| :||||| |||| :|:| ||. |
Qy        5 LLEAEPPHVLISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFDQVRL 63

Db      276 lqsawmeillilqivyrslpyddklayaedyimdeehsrlv-gllelyrailqlvrrykkl 334
            |:| |||:|::|:: ||:      || :| | ::| : :: | |:||:: :|     |:: |
Qy       64 LESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSRFREL 123

Db      335 kvekeefvmlkalalansdsmy--i---enleavqklqdllh---ealq-dyels--qrh 382
            |::  |:: :||: | || |||  :     :: :: :||  ||:    :||          |      :
Qy      124 KLQHKEYLCVKAMILLNS-SMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKSGISSQ 192

Db      384 eeprraqkllltlpllrqtaakavqhfysvklqgkvpmhklflemleakv 433
            ::: | ::||: |: :|::: |::|: ::|      ||:  |:||||:| |
Qy      183 QQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNAHV 232


RESULT    12
ENTRY              S58087      #type complete
TITLE              estrogen receptor related protein - mouse
ORGANISM           #formal_name Mus musculus #common_name house mouse
DATE               13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change
                     26-Apr-1996
ACCESSIONS         S58087
REFERENCE          S58087
   #authors        Pettersson, K., Svensson, K., Mattsson, R., Carlsson, B.;
                     Ohlsson, R.; Berkenstam, A.
   #submission     submitted to the EMBL Data Library, July 1995
   #description    Expression of a novel member of estrogen response
                     element-binding nuclear receptors is restricted to the
                     early stages of chorion formation during mouse
                     embryogenesis.
   #accession      S58087
      ##status          preliminary
      ##molecule_type mRNA
      ##residues        1-433 ##label PET
      ##cross-references EMBL:X89594
CLASSIFICATION     #superfamily erbA transforming protein homology
FEATURE
   101-352              #domain erbA transforming protein homology #label ERBA
SUMMARY            #length 433  #molecular-weight 48400  #checksum 1075

  Query Match              27.2%;  Score 470;  DB 14;  Length 433;
  Best Local Similarity   34.5%;  Pred. No. 5.05e-54;
  Matches    81;  Conservative    68;  Mismatches   70;  Indels   16;  Gaps   10;
```

```
Db      212 lvsellqveqdk-lyamppndipegdikalttlcdladrelvfllnwakhipgfpeltlg 270
             :|  ||  |      | : |:  : |: :: :: |   |||:||| :|:||| |||| |:|
Qy        1 LVLTLLEAEPPHVLISRPSAPFTEASMM-MS-LTKLADKELVHMISWAKKIPGFVELSLF 58

Db      271 dqmsllqsawmeililgivyrslpyddklayaedyimdeehsrlv-glldlyrailqlvr 329
            ||: ||:| |||:|::|:: ||:      || :| | ::| : :: | |:|::: :|
Qy       59 DQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTS 118

Db      330 rykklkvekeefmilkalalansdsmy--i---enleavqklqdllh---ealq-dyels 380
            |:: ||::  |:: :||: | || |||   :     :: :: :||  ||:   :||         |
Qy      119 RFRELKLQHKEYLCVKAMILLNS-SMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKS 177

Db      381 --qrheeprraqkllwtlpllrqtaakavqhfysvklqgkvpmhklflemleakv 433
                :... | ..|| |. .|..: |.::|: ::|      ||:  |:||||:| |
Qy      178 GISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNAHV 232


RESULT    13
ENTRY              A29345      #type complete
TITLE              steroid hormone receptor ERR1 precursor - human
ALTERNATE_NAMES    estrogen-related receptor
ORGANISM           #formal_name Homo sapiens #common_name man
DATE               15-Dec-1988 #sequence_revision 15-Dec-1988 #text_change
                     03-Jun-1996
ACCESSIONS         A29345; A49074
REFERENCE          A29345
   #authors        Giguere, V.; Yang, N.; Segui, P.; Evans, R.M.
   #journal        Nature (1988) 331:91-94
   #title          Identification of a new class of steroid hormone receptors.
   #cross-references MUID:88122546
   #accession      A29345
      ##molecule_type mRNA
      ##residues        1-521 ##label GIG
      ##cross-references EMBL:X51416; EMBL:Y00290
REFERENCE          A49074
   #authors        Wiley, S.R., Kraus, R.J., Zuo, F., Murray, E.E., Loritz, K.;
                     Mertz, J.E.
   #journal        Genes Dev. (1993) 7:2206-2219
   #title          SV40 early-to-late switch involves titration of cellular
                     transcriptional repressors.
   #accession      A49074
      ##status          preliminary
      ##molecule_type protein
      ##residues        166-169,'X',171-173 ##label WIL
CLASSIFICATION     #superfamily erbA transforming protein homology
KEYWORDS           DNA binding; nucleus; zinc finger
FEATURE
   174-434              #domain erbA transforming protein homology #label ERBA\
   176-196              #region zinc finger\
   212-236              #region zinc finger
SUMMARY            #length 521  #molecular-weight 55742  #checksum 5840

  Query Match              25.7%;  Score 444;  DB 13;  Length 521;
  Best Local Similarity   36.4%;  Pred. No. 7.32e-50;
  Matches    84;  Conservative    51;  Mismatches   89;  Indels    7;  Gaps    4;

Db      294 lvshllvvepeklyampdpagpdghlpavatlcdlfdreivvtiswakgipgfsslslsd 353
            ||  ||   ||  :   :: |::|       : : :|  | |:|:|  ||||| ||||  ||| |
Qy        1 LVLTLLEAEPPHVL-ISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 59

Db      354 qmsvlqsvwmevlvlgvaqrslplqdelafaedlvld-eegaraaglgelgaallqlvrr 412
            |: :|:| |||||::|: ||:       | || ||||| :||      |: |:     ||      |
Qy       60 QVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSR 119

Db      413 lqalrlereeyvllkalalansdsvhiedeprlwssceklh-eal---leyeagragp 467
            :: |:|:: ||: :||: | ||      :      :     |   || |   :|:     |      :::|
Qy      120 FRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKSGI 179

Db      468 gggaerrragrllltlpllrqtagkvlahfyqvklegkvpmhklflemlea 518
            ::   :   | : ||: |: :|::: | : |:   :|      ||:  |:||||:|
Qy      180 SSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNA 230
```

```
RESULT    14
ENTRY               I53280      #type complete
TITLE               progesterone receptor - rat
ORGANISM            #formal_name Rattus norvegicus #common_name Norway rat
DATE                02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change
                      02-Aug-1996
ACCESSIONS          I53280
REFERENCE           I53280
   #authors         Park-Sarge, O.K.; Mayo, K.E.
   #journal         Endocrinology (1994) 134:709-718
   #title           Regulation of the progesterone receptor gene by gonadotropins
                      and cyclic adenosine 3',5'-monophosphate in rat granulosa
                      cells.
   #cross-references MUID:94130817
   #accession       I53280
      ##status            preliminary; translated from GB/EMBL/DDBJ
      ##molecule_type mRNA
      ##residues          1-923 ##label RES
      ##cross-references GB:U16922; NID:g463282; CDS_PID:g463283
SUMMARY             #length 923  #molecular-weight 99407  #checksum 9843

  Query Match              19.7%;  Score 340;  DB 14;  Length 923;
  Best Local Similarity   28.5%;  Pred. No. 1.65e-33;
  Matches   59;  Conservative   64;  Mismatches   70;  Indels   14;  Gaps   14;

Db     677 linllmsiepdviyaghdntkpdtsssllts lnqlgerqllsvvkwskslpgfrnlhidd 736
           |:  |:  ||  |   :: ::  | | :|:: ||  |...  |       |:|  |||  | : |
Qy       1 LVLTLLEAEPPHVLISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 59

Db     737 qitliqyswmslmvfglgwrsykhvsgqmlyfapdliln-eq-r-mke-lsfyslcltmw 792
           .: |::  || ::: || |||   | :|  :: |||||:|: :: :  :    | :: : | :
Qy      60 QVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDM-L-L- 114

Db     793 qipqefvklqvtheeflcmkvllllnt-i-pleglrsqsqfeemrssyirelikaig-lr 849
                |   |  :  |  |:||:| ::|||: : ||      :::   : :  :  :  |:  :
Qy     115 ATTSRFFELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVI 174

Db     850 qkgvvpssqrfyqltklldslhdlvkq 876
            |: ::| |:  :|::||  |    |::
Qy     175 AKSGISSQQQSMRLANLL-MLLSHVRH 200


RESULT    15
ENTRY               QRHUP       #type complete
TITLE               progesterone receptor form B - human
ALTERNATE_NAMES     hPR
CONTAINS            progesterone receptor form A
ORGANISM            #formal_name Homo sapiens #common_name man
DATE                30-Jun-1987 #sequence_revision 18-Nov-1994 #text_change
                      15-Nov-1996
ACCESSIONS          S09971; S12464; A03245
REFERENCE           S09971
   #authors         Kastner, P.; Krust, A.; Turcotte, B.; Stropp, U.; Tora, L.;
                      Gronemeyer, H.; Chambon, P.
   #journal         EMBO J. (1990) 9:1603-1614
   #title           Two distinct estrogen-regulated promoters generate
                      transcripts encoding the two functionally different human
                      progesterone receptor forms A and B.
   #cross-references MUID:90228361
   #accession       S09971
      ##molecule_type mRNA
      ##residues          1-933 ##label KAS
      ##cross-references EMBL:X51730
REFERENCE           S12464
   #authors         Kastner, P.
   #submission      submitted to the EMBL Data Library, February 1990
   #accession       S12464
      ##molecule_type mRNA
      ##residues          1-343,'T',345-933 ##label KA2
      ##cross-references EMBL:X51730
REFERENCE           A03245
   #authors         Misrahi, M.; Atger, M.; D'Auriol, L.; Loosfelt, H.; Meriel,
                      C.; Fridlansky, F.; Guiochon-Mantel, A.; Galibert, F.;
                      Milgrom, E.
   #journal         Biochem. Biophys. Res. Commun. (1987) 143:740-748
   #title           Complete amino acid sequence of the human progesterone
                      receptor deduced from cloned cDNA.
   #cross-references MUID:87184565
   #accession       A03245
      ##molecule_type mRNA
      ##residues          1-225,'G',227-255,'V',257-659,'V',661-933 ##label MIS
      ##cross-references GB:M15716
GENETICS
   #gene            GDB:PGR
      ##cross-references GDB:119493
   #map_position 11q22.1-11q22.3
CLASSIFICATION    #superfamily progesterone receptor; erbA transforming protein
                      homology
KEYWORDS          alternative splicing; DNA binding; glycoprotein; nucleus;
                      phosphoprotein; steroid hormone receptor; transcription
                      regulation; zinc finger
FEATURE
   1-933                 #product progesterone receptor form B #status predicted
                           #label MA1\
   165-933               #product progesterone receptor form A #status predicted
                           #label MA2\
   565-629               #domain erbA transforming protein homology #label ERBA\
   567-587               #region zinc finger CCCC motif\
   603-627               #region zinc finger CCCC motif\
   681-933               #domain steroid binding #status predicted #label SIB\
   41                    #binding_site phosphate (Thr) (covalent) #status
                           predicted\
   227,232,552,793       #binding_site phosphate (Ser) (covalent) #status
                           predicted\
   329,374,601           #binding_site phosphate (Tyr) (covalent) #status
                           predicted\
   664                   #binding_site carbohydrate (Asn) (covalent) #status
                           predicted
SUMMARY             #length 933  #molecular-weight 98928  #checksum 5865

  Query Match              19.6%;  Score 338;  DB 3;  Length 933;
  Best Local Similarity   28.8%;  Pred. No. 3.35e-33;
  Matches   59;  Conservative   61;  Mismatches   75;  Indels   10;  Gaps   10;

Db     687 linllmsiepdviyaghdntkpdtsssllts lnqlgerqllsvvkwskslpgfrnlhidd 746
           |:  |    ||  :  :      | |  |     ||  |:::::|: :: |:| :||| :| : |
Qy       1 LVLTLLEAEPPHVLISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 59

Db     747 qitliqyswmslmvfglgwrsykhvsgqmlyfapdlilneqrmk-essfyslcltmwqip 805
           |: |..  || ::. || |||  | .| .. ||||||:|. .  |    ..  :    :
Qy      60 QVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDMLLATT 117

Db     806 qefvklqvsqeeflcmkvllllntipleglrsqtq-feemrs-syi-relikaia-lrqk 861
             |   |  . . :.||:| ::|||: ::  | : ||  :  |  : :   :  |:  :  |
Qy     118 SRFFELKLQHKEYLCVKAMILLNS-SMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAK 176

Db     862 gvvsssqrfyqltklldnlhdlvkq 886
           : :|| |:  :|::||  |   |::
Qy     177 SGISSQQQSMRLANLL-MLLSHVRH 200


Search completed: Thu Dec 18 08:35:12 1997
Job time : 71 secs.
```

```
**********************************************************************
                                                                 (TM)
**********************************************************************

          Release 2 1D John F  Collins  Biocomputing Research Unit
          Copyright (c) 1993  1994, 1995  University of Edinburgh  U K
                 Distribution rights by IntelliGenetics, Inc.

MPsrch_pp   protein - protein database search  using Smith-Waterman algorithm

Run on         Thu Dec 19 08:36:29 1997,  MasPar time 2.12 Seconds
                                           363.743 Million cell updates/sec
Tabular output not generated.

Title:           >US-08-826-361A-4
Description:     (1-233) from US08826361A.pep
Perfect Score:   1726
Sequence:        1 IVLTLLEAEPPHVLISRPSA ........CKNVVPVYDLLLEMLNAHVL 233

Scoring table    PAM 150
                 Gap 11

Searched:        54101 seqs, 4836506 residues

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database:        a-issued
                   1:back1 2:51 3:52 4:53 5:54 6:55 7:56 8:PCT90 9:PCT91
                   10:PCT92 11:PCT93 12:PCT94 13:PCT95 14:PCT96

Statistics:      Mean 29.997,  Variance 146.131;  scale 0.204

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed
      and is derived by analysis of the total score distribution.

                                 SUMMARIES
                    %
Result             Query
  No.   Score     Match Length DB  ID          Description           Pred. No.
-----------------------------------------------------------------------------
    1    1102      63.8    410  1  5223606-5   Patent No. 5223606    5.76e-94
    2     336      19.5    363  1  5223606-6   Patent No. 5223606    1.44e-18
    3     312      18.1    356  1  5223606-7   Patent No. 5223606.   1.34e-16
    4     299      17.3    533  5  US-07-952-  Sequence 2, Applicatio 1.54e-15
    5     296      17.1    446  5  US-07-952-  Sequence 4, Applicatio 2.70e-15
    6     293      17.0    795  3  US-07-716-  Sequence 5, Applicatio 4.75e-15
    7     286      16.6    462  9  PCT-US91-0  Sequence 2, Applicatio 1.76e-14
    8     286      16.6    467  9  PCT-US91-0  Sequence 4, Applicatio 1.76e-14
    9     280      16.2    462  9  PCT-US91-0  Sequence 6, Applicatio 5.29e-14
   10     211      12.2    558  6  US-08-333-  Sequence 12, Applicati 1.75e-08
   11     201      11.6    454  1  5260432-2   Patent No. 5260432.   1.06e-07
   12     191      11.1     60 12  PCT-US94-1  Sequence 1, Applicatio 6.37e-07
   13     190      11.0    368  1  5223606-3   Patent No. 5223606.   7.61e-07
   14     190      11.0    448  1  5223606-2   Patent No. 5223606    7.61e-07
   15     183      10.6    462  1  5171671-2   Patent No. 5171671.   2.64e-06
   16     182      10.5    462 10  PCT-US92-0  Sequence 4, Applicati 3.15e-06
   17     182      10.5    797 10  PCT-US92-0  Sequence 2, Applicatio 3.15e-06
   18     162       9.4    512  7  US-08-464-  Sequence 2, Applicati [illegible]
   19     146       8.5    410  1  5438126-2   Patent No. 5438126.   1.67e-03
   20     141       8.2    272 13  PCT-US95-0  Sequence 23, Applicati 3.91e-03
   21     141       8.2    272  6  US-08-485-  Sequence 23, Applicati 3.91e-03
   22     141       8.2    443  7  US-08-342-  Sequence 4, Applicatio 3.91e-03
   23     139       8.1    446 13  PCT-US95-1  Sequence 3, Applicatio 5.48e-03
   24     138       8.0    461 13  PCT-US95-1  Sequence 2, Applicatio 6.49e-03
   25     138       8.0    461 13  PCT-US95-1  Sequence 2, Applicatio 6.49e-03
   26     138       8.0    461  7  US-08-330   Sequence 2, Applicatio 6.49e-03
   27     138       8.0    461  7  US-08-330-  Sequence 2, Applicati 6.49e-03
   28     135       7.8    [illegible]  7  US-08-342-  Sequence [illegible], Applicatio [illegible]
   29     133       7.7    440  6  US-08-333-  Sequence 8, Applicatio 1.51e-02
   30     131       7.6    433 12  PCT-US94-0  Sequence 2, Applicatio 2.10e-02
   31     [illegible]       7.4    355  1  5223606-4   Patent No. 5223606.   [illegible]
   32     128       7.4    367  3  US-07-727-  Sequence 4, Applicatio 2.47e-02
   33     106       6.1    548  6  US-08-333-  Sequence 2, Applicatio 1.24e+00
   34      99       5.7    451 13  PCT-US95-1  Sequence 2, Applicatio 3.74e+00
   35      99       5.7    484 13  PCT-US95-1  Sequence 1, Applicatio 3.74e+00
   36      97       5.6    298  6  US-08-095-  Sequence 4, Applicatio 5.10e+00
   37      97       5.6    298  6  US-08-096-  Sequence 4, Applicatio 5.10e+00
   38      97       5.6    298  6  US-08-093-  Sequence 4, Applicatio 5.10e+00
   39      97       5.6    307  6  US-08-096-  Sequence 2, Applicatio 5.10e+00
   40      97       5.6    307  6  US-08-095-  Sequence 2, Applicatio 5.10e+00
   41      97       5.6    307  6  US-08-093-  Sequence 2, Applicatio 5.10e+00
   42      96       5.5    462 13  PCT-US95-1  Sequence 2, Applicatio 5.95e+00
   43      91       5.3    302  5  US-07-783-  Sequence 1, Applicatio 1.28e+01
   44      90       5.2    601  6  US-08-333   Sequence 14, Applicati 1.73e+01
   45      88       5.1    478  6  US-08-127   Sequence 11, Applicati [illegible]

                                 ALIGNMENTS

RESULT     1
ID   5223606-5     STANDARD;      PRT;   444 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Patent No. 5223606.
XX
CC   Patent No. 5223606
CC       APPLICANT:  BLAUDIN DE THE, HUGHES;MARCHIO, AGNES;TIOLLAIS,
CC   PIERRE;DEJEAN, ANNE
CC       TITLE OF INVENTION: STEROID/THYROID HORMONE RECEPTOR-RELATED
CC   PROTEIN INAPPROPRIATELY EXPRESSED IN HUMAN HEPATOCELLULAR CARCINOMA
CC       NUMBER OF SEQUENCES: 11
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/134,130
CC         FILING DATE: 17-DEC-1987
CC       PRIOR APPLICATION DATA:
CC   SEQ ID NO:5:
CC         LENGTH: 410
SQ   SEQUENCE   444 AA; 50020 MW; 1110634 CN;

  Query Match             63.8%;  Score 1102;  DB 1;  Length 410;
  Best Local Similarity  59.0%;  Pred. No. 5.76e-94;
  Matches  138;  Conservative   60;  Mismatches  35;  Indels   1;  Gaps   1;

Db      131 MVSALLDAEPILYSEYDPTRPFSEASMMGLLTNLADRELVHMINWAKRVPGFVDLTLHDQ 190
            .: .||.|||        |. ||:||||||  ||:|||:||||||:|||::||||:|:| ||
Qy        1 IVLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFDQ 60

Db      191 VHLLECAWLEILMIGLVWRSMEHPVKLLFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRF 250
            |:|||  |.|:|||.|||:|||..|| ||.|||.|.|||.:|||||||::||||||||||:|||
Qy       61 VRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSRF 120

Db      251 RMMNLQGEEFVCLKSIILLNSGVYTFLSSTLKSLEEKDHIHRVLDKITDTLIHLMAKAGL 310
            | ..||  |..|.|:.|||||..| :...|    . :      : ::|: :||:|: ::||:::
Qy      121 RELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSRKLAHLLNAVTDALVWVIAKSGI 179

Db      311 TLQQQHQRLAQLLLILSHIRHMSNKGMEHLYSMKCKNVVPLYDLLLEMLDAHRL 364
            . |||  ||| ||::|||:|| ||||||||  :|||||||||:|||||||||:|| |
Qy      180 SSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNAHVL 233


RESULT     2
ID   5223606-6     STANDARD;      PRT;   393 AA.
```

```
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Patent No. 5223606.
XX
CC   Patent No. 5223606
CC       APPLICANT:  BLAUDIN DE THE, HUGHES;MARCHIO, AGNES;TIOLLAIS,
CC   PIERRE;DEJEAN, ANNE
CC       TITLE OF INVENTION: STEROID/THYROID HORMONE RECEPTOR-RELATED
CC   PROTEIN INAPPROPRIATELY EXPRESSED IN HUMAN HEPATOCELLULAR CARCINOMA
CC       NUMBER OF SEQUENCES: 11
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/134,130
CC         FILING DATE: 17-DEC-1987
CC       PRIOR APPLICATION DATA:
CC   SEQ ID NO:6:
CC           LENGTH: 363
SQ   SEQUENCE   393 AA; 44638 MW; 907537 CN;

  Query Match             19.5%;  Score 336;  DB 1;  Length 363;
  Best Local Similarity   27.9%;  Pred. No. 1.44e-18;
  Matches   57;  Conservative   62;  Mismatches   77;  Indels    8;  Gaps    8;

Db       117 LINLIMSIFPDVIYAGHDNTKPDTSSSLLTSLNQLGERQLLSVVKWSKSLPGFRNLHIDD 176
             |:  |:  ||  :  :: ::  || :|:: ||  |::::|: :: |:| :||| :| : |
Qy         1 LVLTLLEAEPPHVLISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 59

Db       177 QITLIQYSWMSLMVFGLGWRSYKHVSGQMLYFAPDLILNEQRMK-ESSFYSLCLTMWQIP 235
             |: |::  ||  :::  || |||  |  :|  :: ||||||:|::  |   ::  :   :
Qy        60 QVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDMLLAT 117

Db       236 QEFVKLQVSQEEFLCMKVLLLLNT-I-PLEGLRSQSQFEEMRSSYIRELIKAIG-LRQKG 292
                | | : : |:||:| ::|||: : ||     :::   : :  :  :  |:  :  |:
Qy       118 SRFRELKLQHKEYLCVKAMILLNSSMYPLVTAIQDADSSRKLAHLLNAVTDALVWVIAKS 177

Db       293 VVSSSQRFYQLTKLLDNLHDLVKQ 316
              :|| |:  :|::||  |   |::
Qy       178 GISSQQQSMRLANLL-MLLSHVRH 200


RESULT     3
ID   5223606 7       STANDARD;      PRT;    386 AA
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Patent No. 5223606.
XX
CC   Patent No. 5223606
CC       APPLICANT:  BLAUDIN DE THE, HUGHES;MARCHIO, AGNES;TIOLLAIS,
CC   PIERRE;DEJEAN, ANNE
CC       TITLE OF INVENTION: STEROID/THYROID HORMONE RECEPTOR-RELATED
CC   PROTEIN INAPPROPRIATELY EXPRESSED IN HUMAN HEPATOCELLULAR CARCINOMA
CC       NUMBER OF SEQUENCES: 11
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/134,130
CC         FILING DATE: 17-DEC-1987
CC       PRIOR APPLICATION DATA:
CC   SEQ ID NO:7:
CC           LENGTH: 356
SQ   SEQUENCE   386 AA; 44008 MW; 864027 CN;

  Query Match             18.1%;  Score 312;  DB 1;  Length 356;
  Best Local Similarity   33.8%;  Pred. No. 1.34e-16;
  Matches   49;  Conservative   42;  Mismatches   47;  Indels    7;  Gaps    6;

Db       112 LV-SLLEVIEPEVLYAGTDSSVPDSTWRIMTTLNMLGGRQVIAAVKWAKAIPGFRNLHLD 170
             || :|||   |  || :   ::  :::  :|  :|  :|:  ::::  :  ||| ||||  :| |
Qy         1 LVLTLLEAEPPHVLISRPSAPFTEAS--MMMSLTKLADKELVHMISWAKKIPGFVELSLF 58
```

```
Db       171 DQNTLLQYSWM-FLMAALGWRSYRQSSANLLCFAPDLIIN-EQRMTLPCMYDQCKHMLYV 228
             ||  ||:   || ||  :| |||  :::  :|: |||||:::  :: :  :  :  :    :|
Qy        59 DQVRLLESCWMEVLMMGLMWRSIDHPG-KLI-FAPDLVLDRDEGKCVEGILEIFDMLLAT 116

Db       229 SSELHRLQVSYEEYLCMKYLLLLSS 253
             :| :: | :   ||||:|  ::||:|
Qy       117 TSRFRELKLQHKEYLCVKAMILLNS 141


RESULT     4
ID   US-07-952-800-2      STANDARD;      PRT;    533 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 2, Application US/07952800.
XX
CC    Sequence 2, Application US/07952800
CC    Patent No. 5403925
CC     GENERAL INFORMATION:
CC       APPLICANT:  OZATO, KEIKO
CC       TITLE OF INVENTION:  A NEW MEMBER OF THE NUCLEAR HORMONE
CC       TITLE OF INVENTION:  RECEPTOR SUPERFAMILY AND A CDNA CLONE THEREOF
CC       NUMBER OF SEQUENCES:  4
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  TOWNSEND AND TOWNSEND
CC         STREET:  1 MARKET PLAZA, STEUART TOWER, SUITE 2000
CC         CITY:  SAN FRANCISCO
CC         STATE:  CA
CC         COUNTRY:  USA
CC         ZIP:  94105
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/952,800
CC         FILING DATE:  19920928
CC         CLASSIFICATION:  435
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Weber, Kenneth A
CC         REGISTRATION NUMBER:  31,677
CC         REFERENCE/DOCKET NUMBER:  15290-21-1
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  415-543-9600
CC         TELEFAX:  415-543-5043
CC     INFORMATION FOR SEQ ID NO:  2:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  533 amino acids
CC         TYPE:  AMINO ACID
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   533 AA; 56921 MW; 1406341 CN;

  Query Match             17.3%;  Score 299;  DB 5;  Length 533;
  Best Local Similarity   32.8%;  Pred. No. 1.54e-15;
  Matches   66;  Conservative   52;  Mismatches   70;  Indels   1 ;  Gaps   10;

Db       343 ADKQLFILVEWAKRIPHFSSLPLDDQVILLRAGWNELLIASFSHRSIDVRDGILLAIGLH 402
             |||:|  ::  ||:||  |  |:|  |||  ||  :  | |:|: :: ||||    :::|  '
Qy        35 ADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLV 94

Db       403 VHRNSAHSA-GVGAIFDRVLTELVSKMRDMRMDKTELGCLRAIILFNPDAKGLSNPS-EV 460
             : |: :    |:  ||| :|:    |: |:::::  |  |::|:||:|:    |  :: :
Qy        95 LDRDEGKCVEGILEIFDMLLAT-TSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA 153

Db       461 EV---L-R--EKVYASLE-TYCKQKYP-EQQG-RFAKLLLRLPALRSIGLKCLEHLFFFK 511
             :    |  :  : |  :|     |    : :|||:|:|:||: |: :|  : | :|||:  |
Qy       154 DSSRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMK 213

Db       512 LIGDTPIDTFLMEMLEAPHQL 532
```

```
                  |:   :|:|||:| | |
Qy     214 CKNVVPVYDLLLEMLNA-HVL 233


RESULT    5
ID   US-07-952-800-4     STANDARD;      PRT;   446 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 4, Application US/07952800.
XX
CC    Sequence 4, Application US/07952800
CC    Patent No. 5403925
CC     GENERAL INFORMATION:
CC       APPLICANT:  OZATO, KEIKO
CC       TITLE OF INVENTION:  A NEW MEMBER OF THE NUCLEAR HORMONE
CC       TITLE OF INVENTION:  RECEPTOR SUPERFAMILY AND A CDNA CLONE THEREOF
CC       NUMBER OF SEQUENCES:  4
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  TOWNSEND AND TOWNSEND
CC         STREET:  1 MARKET PLAZA, STEUART TOWER, SUITE 2000
CC         CITY:  SAN FRANCISCO
CC         STATE:  CA
CC         COUNTRY:  USA
CC         ZIP:  94105
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/952,800
CC         FILING DATE:  19920928
CC         CLASSIFICATION:  435
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Weber, Kenneth A
CC         REGISTRATION NUMBER:  31,677
CC         REFERENCE/DOCKET NUMBER:  15280-21-1
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  415-543-9600
CC         TELEFAX:  415-543-5043
CC     INFORMATION FOR SEQ ID NO:  4:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  446 amino acids
CC         TYPE:  AMINO ACID
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   446 AA; 48022 MW; 964946 CN,

  Query Match             17.1%;  Score 296;  DB 5,  Length 446;
  Best Local Similarity   32.8%;  Pred. No. 2.70e-15.
  Matches   66,  Conservative   51,  Mismatches  71,  Indels  13,  Gaps  10,

Db     256 ADKQLFTLVEWAKRIPHFSSLPLDDQVILLRAGWNELLIASFSHRSIDVRDGILLATGLH 315
           ||:|   :: ||:||| |   |:| ||| ||  . | |.|. ..  ||||      ...|  |
Qy      35 ADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLV 94

Db     316 VHRNSAHSA-GVGAIFDRVLTELVSKMRDMRMDKIELGCLRAIILFNPDAKGL-SNPGEV 373
           : |. .     ::  |||  .||.    |   |           :   :  |.||  :.      |
Qy      95 LDRDEGKCVEGILEIFDMLLAT-TSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA 153

Db     374 EI---L-R--EKVYASLE-TYCKQKYP-EQQG-RFAKLLLRLPALRSIGLKCLEHLFFFK 424
           :     | :  : |  :|       |      . .||: |.|.||. |. :|   : | :|||:  |
Qy     154 DSSFKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMK 213

Db     425 LIGDTPIDTFLMEMLEAPHQL 445
                |:  :|:|||:| | |
Qy     214 CKNVVPVYDLLLEMLNA-HVL 233


RESULT    6
```

```
ID   US-07-716-827C-5     STANDARD;      PRT;   795 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 5, Application US/07716827C.
XX
CC    Sequence 5, Application US/07716827C
CC    Patent No. 5215916
CC     GENERAL INFORMATION:
CC       APPLICANT:  Simons Jr., Stoney S.
CC       APPLICANT:  Yamamoto, K. R.
CC       APPLICANT:  Chakraborti, P. K.
CC       APPLICANT:  Garabedian, M. J.
CC       TITLE OF INVENTION:  SUPER GLUCOCORTICOID RECEPTORS
CC       NUMBER OF SEQUENCES:  5
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Cushman, Darby & Cushman
CC         STREET:  Eleventh floor, 1615 L Street, N.W.
CC         CITY.  Washington
CC         STATE:  D.C.
CC         COUNTRY:  U.S.A.
CC         ZIP:  20036-5601
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM.  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA.
CC         APPLICATION NUMBER:  US/07/716,827C
CC         FILING DATE:  19910619
CC         CLASSIFICATION:  435
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Scott, Watson T.
CC         REGISTRATION NUMBER:  26,581
CC         REFERENCE/DOCKET NUMBER:  WTS/5683/94453
CC       TELECOMMUNICATION INFORMATION.
CC         TELEPHONE:  (202)8613000
CC         TELEFAX:  (202)822-8944
CC         TELEX:  6714627 CUSH
CC     INFORMATION FOR SEQ ID NO:  5:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  795 amino acids
CC         TYPE:  AMINO ACID
CC         TOPOLOGY:  linear
SQ   SEQUENCE   795 AA; 87498 MW; 3326890 CN;

  Query Match             17.0%,  Score 293,  DB 3,  Length 795;
  Best Local Similarity   31.7%,  Pred. No. 4.75e-15;
  Matches   46;  Conservative   46;  Mismatches  47;  Indels   6;  Gaps   5;

Db     550 LV-SLLEVIEPEVLYAGYDSSVPDSAWRIMTTLNMLGGRQVIAAVKWAKAIPGLRNLHLD 608
           || :|||   | || :   ::  :::  :| :| :|: ::::   : ||| | |: :|  '
Qy       1 LVLTLLEAEPPHVLISRPSAPFTEAS--MMMSLTKLADKELVHMISWAKKIPGFVELSLF 58

Db     609 DQMTLLQYSWMFLMAFALGWRSYRQSSGNLLCFAPDLIIN-EQRMSLPCMYDQCKHMLFV 667
           ||: ||:  || ::  :| |||   :: |:|: ||||:::  :: :  :  :  :     :|
Qy      59 DQVRLLESCWMEVLMMGLMWRSIDHP-GKLI-FAPDLVLDRDEGKCVEGILEIFDMLLAT 116

Db     668 SSELQRLQVSYEEYLCMKTLLLLSS 692
           :| :: | :    ||||:|:::||:|
Qy     117 TSRFRELKLQHKEYLCVKAMILLNS 141


RESULT    7
ID   PCT-US91-00399-2     STANDARD;      PRT;   462 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 2, Application PC/TUS9100399.
```

```
XX
CC    Sequence 2, Application PC/TUS9100399
CC     GENERAL INFORMATION:
CC       APPLICANT:  Mangelsdorf, Dr., David J.
CC       APPLICANT:  Evans Dr., Ronald M.
CC       TITLE OF INVENTION:  RETINOID RECEPTOR COMPOSITIONS AND
CC       TITLE OF INVENTION:  METHODS
CC       NUMBER OF SEQUENCES:  7
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  FITCH, EVEN, TABIN & FLANNERY
CC         STREET:  135 South LaSalle Street, Suite 900
CC         CITY:  Chicago
CC         STATE:  Illinois
CC         COUNTRY:  USA
CC         ZIP:  60603
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US91/00399
CC         FILING DATE:  19910122
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/478,071
CC         FILING DATE:  09-FEB-1990
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Reiter, Mr., Stephen E.
CC         REGISTRATION NUMBER:  31192
CC         REFERENCE/DOCKET NUMBER:  50852
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (619) 552 1311
CC         TELEFAX:  (619) 552 0095
CC         TELEX:  20 6566 FATLAW CGO
CC     INFORMATION FOR SEQ ID NO:  2:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  462 amino acids
CC         TYPE:  AMINO ACID
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   462 AA;  50811 MW;  1063517 CN;

  Query Match             16.6%;  Score 286;  DB 9;  Length 462;
  Best Local Similarity   31.5%;  Pred. No. 1.76e-14;
  Matches   62;  Conservative   53;  Mismatches   70;  Indels   12;  Gaps    9;

Db        272 ADKQLFTLVEWAKRIPHFSELPLDDQVILLRAGWNELLIASFSHRSIAVKDGILLATGLH 331
              |||:|  :: |||:|| | ||:| ||| || : | |:|. ..   |||       :;|  |
Qy         35 ADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLV 94

Db        332 VHRNSAHSA-GVGAIFDRVLTELVSKMRDMQMDKTELGCLRAIVLFNPDSKGLSNPA-EV 389
              : |: :     |:  ||| :|:     |: |:: ::  |  |::::|:|:     |  :: .
Qy         95 LDRDEGKCVEGILEIFDMLLAT-TSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA 153

Db        390 EA---L-R--EKVYASLE-AYCKHKYP-EQPG-RFAKLLLRLPALRSIGLKCLEHLFFFK 440
              ::   | :  : |  :|        |     : :| : |:|:||: |: :|   : | :|||:  |
Qy        154 DSSRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMK 213

Db        441 LIGDTPIDTFLMEMLEA 457
                   |:   ::|||:|
Qy        214 CKNVVPVYDLLLEMLNA 230


RESULT     8
ID   PCT-US91-00399-4      STANDARD;      PRT;    467 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 4, Application PC/TUS9100399.
XX
```

```
CC    Sequence 4, Application PC/TUS9100399
CC     GENERAL INFORMATION:
CC       APPLICANT:  Mangelsdorf, Dr., David J.
CC       APPLICANT:  Evans Dr., Ronald M.
CC       TITLE OF INVENTION:  RETINOID RECEPTOR COMPOSITIONS AND
CC       TITLE OF INVENTION:  METHODS
CC       NUMBER OF SEQUENCES:  7
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  FITCH, EVEN, TABIN & FLANNERY
CC         STREET:  135 South LaSalle Street, Suite 900
CC         CITY:  Chicago
CC         STATE:  Illinois
CC         COUNTRY:  USA
CC         ZIP:  60603
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US91/00399
CC         FILING DATE:  19910122
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/478,071
CC         FILING DATE:  09-FEB-1990
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Reiter, Mr., Stephen E.
CC         REGISTRATION NUMBER:  31192
CC         REFERENCE/DOCKET NUMBER:  50852
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (619) 552-1311
CC         TELEFAX:  (619) 552-0095
CC         TELEX:  20 6566 PATLAW CGO
CC     INFORMATION FOR SEQ ID NO:  4:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  467 amino acids
CC         TYPE:  AMINO ACID
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   467 AA;  51216 MW;  1083042 CN;

  Query Match             16.6%;  Score 286;  DB 9;  Length 467;
  Best Local Similarity   31.5%;  Pred. No. 1.76e-14;
  Matches   62;  Conservative   53;  Mismatches   70;  Indels   12;  Gaps    9;

Db        277 ADKQLFTLVEWAKRIPHFSELPLDDQVILLRAGWNELLIASFSHRSIAVKDGILLATGLH 336
              |||:|  :: |||:|| | ||:| ||| || : | |:|: ::   |||       :::|  ,
Qy         35 ADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLV 94

Db        337 VHRNSAHSA-GVGAIFDRVLTELVSKMRDMQMDKTELGCLRAIVLFNPDSKGLSNPA-EV 394
              : |: :     |:  ||| :|:     |: |:: ::  |  |::::|:|:     |  :: :
Qy         95 LDRDEGKCVEGILEIFDMLLAT-TSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA 153

Db        395 EA---L-R--EKVYASLE-AYCKHKYP-EQPG-RFAKLLLRLPALRSIGLKCLEHLFFFK 445
              ..    | .   . |  .|        |     . .| . |.|.||. |. .|   . | :|||:  |
Qy        154 DSSRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMK 213

Db        446 LIGDTPIDTFLMEMLEA 462
                   |:   :|:|||:|
Qy        214 CKNVVPVYDLLLEMLNA 230


RESULT     9
ID   PCT-US91-00399-6      STANDARD;      PRT;    463 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 6, Application PC/TUS9100399.
XX
CC    Sequence 6, Application PC/TUS9100399
CC     GENERAL INFORMATION:
CC       APPLICANT: Mangelsdorf, Dr., David J.
CC       APPLICANT: Evans Dr., Ronald M.
CC       TITLE OF INVENTION: RETINOID RECEPTOR COMPOSITIONS AND
CC       TITLE OF INVENTION: METHODS
CC       NUMBER OF SEQUENCES: 7
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE: FITCH, EVEN, TABIN & FLANNERY
CC         STREET: 135 South LaSalle Street, Suite 900
CC         CITY: Chicago
CC         STATE: Illinois
CC         COUNTRY: USA
CC         ZIP: 60603
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE: Floppy disk
CC         COMPUTER: IBM PC compatible
CC         OPERATING SYSTEM: PC-DOS/MS-DOS
CC         SOFTWARE: PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER: PCT/US91/00399
CC         FILING DATE: 19910122
CC         CLASSIFICATION: 435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER: 07/478,071
CC         FILING DATE: 09-FEB-1990
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME: Reiter, Mr., Stephen E.
CC         REGISTRATION NUMBER: 31192
CC         REFERENCE/DOCKET NUMBER: 50852
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE: (619) 552-1311
CC         TELEFAX: (619) 552-0095
CC         TELEX: 20 6566 PATLAW CGO
CC     INFORMATION FOR SEQ ID NO: 6:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH: 463 amino acids
CC         TYPE: AMINO ACID
CC         TOPOLOGY: linear
CC       MOLECULE TYPE: protein
SQ   SEQUENCE   463 AA; 50978 MW; 1097932 CN;

  Query Match             16.2%;  Score 280;  DB 9;  Length 463;
  Best Local Similarity   29.4%;  Pred. No. 5.39e-14;
  Matches   58;  Conservative   57;  Mismatches   70;  Indels   12;  Gaps    9;

Db      273 ADKQIFTLVEWAKRIPHFSDLTLEDQVILLPAGWNELLIASFSHRSVSVQDGILLATGLH 332
            |:|:|   :: :||:||  | :|:| ||| || : | |:|. ..   ||.     ...|   |
Qy       35 ADKFLVHMISWAKKIPGFVELSIFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLV 94

Db      333 VHFSSAHS-FGVGSIFDPVITFLVSFMKDMQMDKSEIGCLRAIVLFNPDAKGLSNPS-EV 390
            : |  :     |:  ||| :|:     |: ... ..  |   |..|.:|:|.      |    .
Qy       95 LDRDFGKCVEGILEIFDMILAT-TSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA 153

Db      391 ET---L-R--EKVYATLE-AYTKQKYP-EQPG-RFAKLLLRLPALRSIGLKCLEHLFFFK 441
            ::   | :  : |  :|      :|     . .| . |.|.||. |. .|   . | .|||.   |
Qy      154 DSSRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMK 213

Db      442 LIGDTPIDSFLMEMLET 458
                |:  :::|||::
Qy      214 CKNVVPVYDLLLEMLNA 230


RESULT   10
ID   US-08-333-358-12     STANDARD;     PRT;   558 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 12, Application US/08333358.
XX
CC    Sequence 12, Application US/08333358
CC    Patent No. 5571696
```

```
CC     GENERAL INFORMATION:
CC       APPLICANT: EVANS Ph.D., RONALD M.
CC       APPLICANT: MANGELSDORF Ph.D., DAVID J.
CC       APPLICANT: ONG Ms., ESTELITA S.
CC       APPLICANT: ORO Ph.D., ANTHONY E.
CC       APPLICANT: BORGMEYER Ph.D., UWE K.
CC       APPLICANT: GIGUERE Ph.D., VINCENT NMN
CC       APPLICANT: YAO Mr., TSO-PANG NMN
CC       TITLE OF INVENTION: NOVEL RECEPTORS
CC       NUMBER OF SEQUENCES: 14
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
CC         STREET: 444 So. Flower St., Suite 2000
CC         CITY: Los Angeles
CC         STATE: CA
CC         COUNTRY: US
CC         ZIP: 90071-2921
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE: Floppy disk
CC         COMPUTER: IBM PC compatible
CC         OPERATING SYSTEM: PC-DOS/MS-DOS
CC         SOFTWARE: PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER: US/08/333,358
CC         FILING DATE:
CC         CLASSIFICATION: 435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER: US/07/761,068
CC         FILING DATE:
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME: Reiter Ph.D., Stephen E.
CC         REGISTRATION NUMBER: 31192
CC         REFERENCE/DOCKET NUMBER: P31 8936
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE: (619) 535-9001
CC         TELEFAX: (619) 535-8949
CC     INFORMATION FOR SEQ ID NO: 12:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH: 558 amino acids
CC         TYPE: amino acid
CC         TOPOLOGY: linear
CC       MOLECULE TYPE: protein
SQ   SEQUENCE   558 AA; 63296 MW; 1624056 CN;

  Query Match             12.2%;  Score 211;  DB 6;  Length 558;
  Best Local Similarity   31.5%;  Pred. No. 1.75e-08;
  Matches   47;  Conservative   36;  Mismatches   62;  Indels    4;  Gaps    4;

Db      341 LEPLGTPMLIE-DGYAVTQAFLFAILCRLADELLFRQIAWIKKLPFFCELSIKDYTCLLS 399
            ||.    .'|     . . |.| .      | .|||  | . |:| ||:| | |||: |    ||
Qy        6 LEAFPPHVLISFPSAFFIEASMMMSLIKLADKELVHMISWAKKIPGFVELSLFDQVRLLE 65

Db      400 STWQELTILLSSLTVYSKQIFGELADVIAKYSPSDEELHRFSDEGMEVIERLIYLYHKFHQ 459
            | | | .:. .|      | .   | |     ..       .|   :   :  :|::: |:    :':: 
Qy       66 SCWMF-VIMMGLMWRSIDHPGKLI-FAPDLVLDRDE-GKCVEGILEIFDMLLAIISRFRE 122

Db      460 LKVSNEEYACMKAINFLNQDIRGLTSASQ 488
            ||: : || |:||: :||   :   | :|:|
Qy      123 LKLQHKEYLCVKAMILLNSSMYPLVTATQ 151


RESULT   11
ID   5260432-2     STANDARD;     PRT;   492 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Patent No. 5260432.
XX
CC   Patent No. 5260432
CC      APPLICANT: TAKAKU, FUMIMARO;ISHIKAWA, TAKASHI;IMAWARI, MICHIO;
CC   EVANS, RONALD M.;UMESONO, KAZUHIKO
CC        TITLE OF INVENTION: HUMAN GAMMA RETINOIC ACID RECEPTOR DNA
CC        NUMBER OF SEQUENCES: 4
CC        CURRENT APPLICATION DATA:
CC          APPLICATION NUMBER: US/07/370,407
CC          FILING DATE: 22-JUN-1989
CC   SEQ ID NO:2:
CC          LENGTH: 454
SQ   SEQUENCE   492 AA; 54527 MW; 1333563 CN;

  Query Match             11.6%;  Score 201;  DB 1;  Length 454;
  Best Local Similarity  26.0%;  Pred. No. 1.06e-07;
  Matches   52;  Conservative   59;  Mismatches  76;  Indels  13;  Gaps   9;

Db     230 FSELATKCIIKIVEFAKRLPGFTGLSIADQITLLKAAC-LDILMLRICTRYTPEQDTMTF 288
            :: || |  :: ::   ||::|||  ||: ||: ||  :| :::||: :  |         . |
Qy      31 LTKLADKELVHMISWAKKIPGFVELSLFDQVRLL-ESCWMEVLMMGLMWRSIDHPGKLIF 89

Db     289 SDGLTLNRTQ-MHNAGFGPLTDLVFAFAGQLLPLEMDDTE----TG--LL-SAICLICGD 340
            :  | |:|  :  :     |:   : |:::| ::::  | ::  |      :  || |::  :
Qy      90 APDLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTA 149

Db     341 RMDLEEPEKVDKLQEPLLEAL-RLYARRRRPSQP-YM-FPRMLMKITDLRGISTKGAERA 397
              | :  : |:  | ::: :|| :: |:    :||    | :: :|| :: :|  | || |:
Qy     150 TQDADSSRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGMEHL 209

Db     398 ITLKMEIPGPMPPLIREMLE 417
            : :|       |:  |: |||:
Qy     210 LNMKCKNVVPVYDLLLEMLN 229


RESULT   12
ID   PCT-US94-14074-1     STANDARD;      PRT;    60 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 1, Application PC/TUS9414074.
XX
CC    Sequence 1, Application PC/TUS9414074
CC     GENERAL INFORMATION:
CC        APPLICANT:  Maksymowych, Andrew B
CC        APPLICANT:  Hsu, Tin-Chen
CC        APPLICANT:  Jameson, Bradford A
CC        APPLICANT:  Litwack, Gerald
CC        TITLE OF INVENTION:  Biologically Active Compounds
CC        NUMBER OF SEQUENCES:  1
CC        CORRESPONDENCE ADDRESS:
CC          ADDRESSEE:  Woodcock Washburn Kurtz MAckiewicz & Norris
CC          STREET:  One Liberty Place 46th Floor
CC          CITY:  Philadelphia
CC          STATE:  Pennsylvania
CC          COUNTRY:  United States of America
CC          ZIP:  19103
CC        COMPUTER READABLE FORM:
CC          MEDIUM TYPE:  Floppy disk
CC          COMPUTER:  IBM PC compatible
CC          OPERATING SYSTEM:  PC-DOS/MS-DOS
CC          SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC        CURRENT APPLICATION DATA:
CC          APPLICATION NUMBER:  PCT/US94/14074
CC          FILING DATE:
CC          CLASSIFICATION:
CC        PRIOR APPLICATION DATA:
CC          APPLICATION NUMBER:  US/08/164,102
CC          FILING DATE:  07-DEC-1993
CC          CLASSIFICATION:
CC        ATTORNEY/AGENT INFORMATION:
CC          NAME:  DeLuca, Mark
CC          REGISTRATION NUMBER:  33,229
CC          REFERENCE/DOCKET NUMBER:  TJU-1418
CC        TELECOMMUNICATION INFORMATION:
CC          TELEPHONE:  215-568-3100
CC          TELEFAX:  215-568-3439
CC      INFORMATION FOR SEQ ID NO.  1:
CC        SEQUENCE CHARACTERISTICS:
CC          LENGTH:  60 amino acids
CC          TYPE:  amino acid
CC          TOPOLOGY:  linear
CC        MOLECULE TYPE:  peptide
SQ   SEQUENCE   60 AA; 6957 MW; 18671 CN;

  Query Match             11.1%;  Score 191;  DB 12;  Length 60;
  Best Local Similarity  40.7%;  Pred. No. 6.37e-07;
  Matches   22;  Conservative   15;  Mismatches  17;  Indels   0;  Gaps   0;

Db       7 IMTTLNMLGGPQVIAAVKWAKAIPGFRNLHLDDQMILLQYSWMFLMAFALGWRS 60
            :| :|  :|:  ....   . ||| |||| .: : ||. ||.   || ::   :| |||
Qy      27 MMMSLTKLADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRS 80


RESULT   13
ID   5223606-3     STANDARD;      PRT;   398 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Patent No. 5223606.
XX
CC   Patent No. 5223606
CC        APPLICANT:  BLAUDIN DE THE, HUGHES;MARCHIO, AGNES;TIOLLAIS,
CC   PIERRE;DEJEAN, ANNE
CC        TITLE OF INVENTION: STEROID/THYROID HORMONE RECEPTOR-RELATED
CC   PROTEIN INAPPROPRIATELY EXPRESSED IN HUMAN HEPATOCELLULAR CARCINOMA
CC        NUMBER OF SEQUENCES: 11
CC        CURRENT APPLICATION DATA:
CC          APPLICATION NUMBER:  US/07/134,130
CC          FILING DATE: 17-DEC-1987
CC        PRIOR APPLICATION DATA:
CC   SEQ ID NO:3:
CC          LENGTH: 368
SQ   SEQUENCE   398 AA; 44842 MW; 907111 CN;

  Query Match             11.0%;  Score 190;  DB 1;  Length 368;
  Best Local Similarity  23.0%;  Pred. No. 7.61e-07;
  Matches   46;  Conservative   66;  Mismatches  75;  Indels  13;  Gaps   9;

Db     141 FSELATKCIIKIVEFAKRLPGFTGLTIADQITLLKAAC-LDILILRICTRYTPEQDTMTF 199
            .: || |  :: ::   ||::|||  |:: ||: ||  :| :::|:: :  |         : |
Qy      31 LTKLADKELVHMISWAKKIPGFVELSLFDQVRLL-ESCWMEVLMMGLMWRSIDHPGKLIF 89

Db     200 SDGLTLNRTQ-MHNAGFGPLTDLVFTFANQLLPLEMDDTE----TG--LL-SAICLICGD 251
            :  | |:|  :  :     |:   : |:::: ::::  | ::  |      :  || |.:  :
Qy      90 APDLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTA 149

Db     252 RQDLEEPTKVDKLQEPLLEALK-IYIRKRRPSKPH-M-FPKILMKITDLRSISAKGAERV 308
             || . . |.  | ::: :||  :   :    :|  : | ::::||  .: :|  | || |::
Qy     150 TQDADSSRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGMEHL 209

Db     309 ITLKMEIPGSMPPLIQEMME 328
            : :|      :: |: ||::
Qy     210 LNMKCKNVVPVYDLLLEMLN 229


RESULT   14
ID   5223606-2     STANDARD;      PRT;   485 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Patent No. 5223606.
XX
CC   Patent No. 5223606
CC        APPLICANT:  BLAUDIN DE THE, HUGHES;MARCHIO, AGNES;TIOLLAIS,
CC   PIERRE;DEJEAN, ANNE
CC      TITLE OF INVENTION: STEROID/THYROID HORMONE RECEPTOR-RELATED
CC   PROTEIN INAPPROPRIATELY EXPRESSED IN HUMAN HEPATOCELLULAR CARCINOMA
CC      NUMBER OF SEQUENCES: 11
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER: US/07/134,130
CC        FILING DATE: 17-DEC-1987
CC      PRIOR APPLICATION DATA:
CC   SEQ ID NO:2:
CC        LENGTH: 448
SQ   SEQUENCE   485 AA; 54437 MW; 1345912 CN;

  Query Match             11.0%;  Score 190;  DB 1;  Length 448;
  Best Local Similarity   23.0%;  Pred. No. 7.61e-07;
  Matches   46;  Conservative   66;  Mismatches   75;  Indels   13;  Gaps    9;

Db      221 FSELATKCIIKIVEFAKRLPGFTGLTIADQITLLKAAC-LDILILRICTRYTPEQDTMTF 279
            :: || | :: ::  ||::|||  |:: ||. ||  :| ...|.. :  |        . |
Qy       31 LTKLADKELVHMISWAKKIPGFVELSLFDQVRLL-ESCWMEVLMMGLMWRSIDHPGKLIF 89

Db      280 SDGLILNRIQ-MHNAGFGPLTDLVFTFANQLLPLEMDDTE----TG--LL-SAICLICGD 331
            :  | |:|  . .    |.  . |.... ....  | :.  |        . || |:: :
Qy       90 APDLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTA 149

Db      332 RQDLEEPTKVDKLQEPLLEALK-IYIRKRRPSKPH-M-FPKILMKITDLRSISAKGAERV 388
             || : : |:  | ::: :||  .    .   .|  . | :...|| :: :|  | || |::
Qy      150 TQDADSSRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGMEHL 209

Db      389 ITLKMEIPGSMPPLIQEMME 408
            : :|      ::  |: ||::
Qy      210 LNMKCKNVVPVYDLLLEMLN 229


RESULT   15
ID   5171671-2     STANDARD;     PRT;    501 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Patent No. 5171671.
XX
CC   Patent No. 5171671
CC      APPLICANT: EVANS, RONALD M.;ONG, ESTELITA S.;SEGUI,
CC   PRUDIMAR S.;THOMPSON, CATHERINE C.;UEMSONO, KAZUHIKO
CC   GIGUERE, VINCENT
CC      TITLE OF INVENTION: RETINOIC ACID RECEPTOR COMPOSITION
CC      NUMBER OF SEQUENCES: 2
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER: US/07/546,256
CC        FILING DATE: 06-AUG-1990
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER: 276,536
CC        FILING DATE: 30-NOV-1988
CC        APPLICATION NUMBER: 128,331
CC        FILING DATE: 02-DEC-1987
CC   SEQ ID NO:2:
CC        LENGTH: 462
SQ   SEQUENCE   501 AA; 55050 MW; 1419525 CN;

  Query Match             10.6%;  Score 183;  DB 1;  Length 462;
  Best Local Similarity   23.5%;  Pred. No. 2.64e-06;
  Matches   47;  Conservative   66;  Mismatches   74;  Indels   13;  Gaps   11;

Db      228 FSELSTKCIIKTVEFAKQLPGFTTLTIADQITLLKAAC-LDILILRICTRYTPEQDTMTF 286
            :: |: | ::  :  |' :'||  |:: ||: |'  :| :::|:: :  | :        : |
Qy       31 LTKLADKELVHMISWAKKIPGFVELSLFDQVRLL-ESCWMEVLMMGLMWRSIDHPGKLIF 89

Db      287 SDGLTLNRTQ-MHNAGFGPLTDLVFAFANQL--LPLE-MDD--AETG-LL-SAICLICGD 338
            :  | |:| : :    |.  . |..:| ...:  | |: ::        :  || |:: :
Qy       90 APDLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTA 149

Db      339 RQDLEQPDRVDMLQEPLLEALK-VYVRKRRPSRPH-M-FPKMLMKITDLRSISAKGAERV 395
             || . . .. . | ... .:| | .   .   :|: . | ....|| .. .|  | || |::
Qy      150 TQDADSSRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGMEHL 209

Db      396 ITLKMEIPGSMPPLIQEMLE 415
            : :|      ::  |: |||:
Qy      210 LNMKCKNVVPVYDLLLEMLN 229

Search completed: Thu Dec 18 06:36:44 1997
Job time : 16 secs.
```

```
*******************************************************************

                  M P S R C H                            (TM)

*******************************************************************

         Release 2.1D John F. Collins, Biocomputing Research Unit
         Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
                Distribution rights by IntelliGenetics, Inc.

MPsrch_nn   n.a. - n.a. database search, using Smith-Waterman algorithm

Run on:          Thu Dec 18 18:21:39 1997;  MasPar time 71.55 Seconds
                                            866.632 Million cell updates/sec
Tabular output not generated

Title:           >US-08-826-361A-1
Description:     (1-1434) from US08826361A seq
Perfect Score:   1434
N.A. Sequence:       1 ATGAATTACAGCATTCCCAG ...... AGAACCCACAGTCTCAGTGA 1434
    Comp.              TACTTAATGTCGTAAGGGTC ...... TCTTGGGTGTCAGAGTCACT

Scoring table:   TABLE default
                 Gap 6

Nmatch   STD :   Dbase 0;  Query 0

Searched:        81563 seqs  21620618 bases x 2

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database:        n-issued
                   1:backl 2:51 3:52 4:53 5:54 6:55 7:56 8:PCT90 9:PCT91
                   10:PCT92 11:PCT93 12:PCT94 13:PCT95 14:PCT96

Statistics:      Mean 8.900;  Variance 4.521;  scale 1.969

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES

                  %
Result          Query
   No.  Score   Match Length DB  ID            Description            Pred. No.
------------------------------------------------------------------------------
     1    197    13.7   4963  5  US-08-076-    Sequence 16, Applicati  4.16e-135
     2    197    13.7   4963  7  US-08-260-    Sequence 9, Applicatio  4.16e-135
c    3     71     5.0   3460  6  US-08-312-    Sequence 1, Applicatio  1.13e-34
     4     63     4.4   6244  5  US-08-076-    Sequence 15, Applicati  1.10e-28
     5     63     4.4   6244  7  US-08-260-    Sequence 8, Applicatio  1.10e-28
c    6     62     4.3   7219  7  US-08-232-    Sequence 14, Applicati  6.06e-28
     7     55     3.8   2989  1  5223606-1     Patent No. 5223606.     8.23e-23
     8     53     3.7   3014  7  US-08-629-    Sequence 1, Applicatio  2.30e-21
     9     51     3.6   1576  1  5260432-1     Patent No. 5260432.     6.30e-20
    10     47     3.3    387  5  US-08-087-    Sequence 1, Applicatio  4.37e-17
    11     47     3.3    387  6  US-08-561-    Sequence 1, Applicatio  4.37e-17
    12     42     2.9   2424 10  PCT-US92-0    Sequence 4, Applicatio  1.32e-13
    13     42     2.9   2940  1  5171671-1     Patent No. 5171671.     1.32e-13
    14     42     2.9   3026 10  PCT-US92-0    Sequence 1, Applicatio  1.32e-13
    15     38     2.6   1662  9  PCT-US91-0    Sequence 5, Applicatio  6.79e-11
    16     35     2.4    215  6  US-08-238-    Sequence 5, Applicatio  5.56e-09
    17     33     2.3   1952  6  US-08-333-    Sequence 1, Applicatio  1.30e-07
c   18     30     2.1    215  6  US-08-238-    Sequence 5, Applicatio  1.02e-05
    19     30     2.1   2095  9  PCT-US91-0    Sequence 3, Applicatio  1.02e-05
    20     30     2.1   2130  5  US-07-952-    Sequence 1, Applicatio  1.02e-05
    21     29     2.0   1866  9  PCT-US91-0    Sequence 1, Applicatio  4.21e-05
    22     28     2.0   2204  5  US-07-952-    Sequence 3, Applicatio  1.71e-04
    23     28     2.0   2219  9  PCT-US91-0    Sequence 7, Applicatio  1.71e-04
    24     28     2.0   2468  6  US-08-333-    Sequence 11, Applicati  1.71e-04
    25     28     2.0   7219  7  US-08-232-    Sequence 14, Applicati  1.71e-04
    26     26     1.8   1649 12  PCT-US94-0    Sequence 1, Applicatio  2.66e-03
    27     26     1.8   1059  6  US-08-333-    Sequence 7, Applicatio  2.66e-03
    28     26     1.8   2304  7  US-08-464-    Sequence 1, Applicatio  2.66e-03
c   29     24     1.7     69  7  US-08-471-    Sequence 142, Applicat  3.78e-02
    30     24     1.7     74 13  PCT-US95-1    Sequence 100, Applicat  3.78e-02
    31     24     1.7     74 13  PCT-US95-1    Sequence 94, Applicati  3.78e-02
c   32     25     1.7     74 13  PCT-US95-1    Sequence 94, Applicati  1.02e-02
c   33     24     1.7     74 13  PCT-US95-1    Sequence 100, Applicat  3.78e-02
    34     25     1.7     75 13  PCT-US95-1    Sequence 99, Applicati  1.02e-02
c   35     24     1.7     75 13  PCT-US95-1    Sequence 99, Applicati  3.78e-02
    36     24     1.7     81 13  PCT-US95-1    Sequence 92, Applicati  3.78e-02
    37     24     1.7     81 13  PCT-US95-1    Sequence 98, Applicati  3.78e-02
c   38     25     1.7     81 13  PCT-US95-1    Sequence 92, Applicati  1.02e-02
c   39     24     1.7     81 13  PCT-US95-1    Sequence 98, Applicati  3.78e-02
    40     25     1.7     92 13  PCT-US95-1    Sequence 97, Applicati  1.02e-02
c   41     25     1.7     92 13  PCT-US95-1    Sequence 97, Applicati  1.02e-02
    42     24     1.7    105  5  US-07-865-    Sequence 13, Applicati  3.78e-02
    43     24     1.7   2009  6  US-08-333-    Sequence 3, Applicatio  3.78e-02
c   44     23     1.6     66  7  US-08-471-    Sequence 144, Applicat  1.37e-01
c   45     23     1.6    105  5  US-07-865-    Sequence 13, Applicati  1.37e-01

                              ALIGNMENTS

RESULT    1
ID   US-08-076-726-16 STANDARD; DNA; UNC; 4963 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 16, Application US/08076726.
CC   Sequence 16, Application US/08076726
CC   Patent No. 5464758
CC    GENERAL INFORMATION:
CC       APPLICANT:  Gossen, Manfred
CC       APPLICANT:  Bujard, Hermann
CC       TITLE OF INVENTION:  Tight Control of Gene Expression in
CC       TITLE OF INVENTION:  Eucaryotic Cells by Tetracycline-responsive Promot
 ers
CC       NUMBER OF SEQUENCES:  16
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Sterne, Kessler, Goldstein & Fox
CC         STREET:  1100 New York Avenue, N.W.
CC         CITY:  Washington
CC         STATE:  District of Columbia
CC         COUNTRY:  United States of America
CC         ZIP:  20005-3934
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/076,726
CC         FILING DATE:  14-JUN-1993
CC         CLASSIFICATION:  435
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Esmond, Robert W.
CC         REGISTRATION NUMBER:  32,893
CC         REFERENCE/DOCKET NUMBER:  1941.143.001
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (202)371-2600
CC         TELEFAX:  (202)371-2540
CC         TELEX:  248636 SSK
CC     INFORMATION FOR SEQ ID NO:  16:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  4963 base pairs
CC         TYPE:  nucleic acid
CC         STRANDEDNESS:  both
CC          TOPOLOGY:  both
SQ   Sequence 4963 BP; 1259 A; 1300 C; 1278 G; 1126 T; 0 other;

  Query Match             13.7%;  Score 197;  DB 5;  Length 4963;
  Best Local Similarity   66.7%;  Pred. No. 4.16e-135;
  Matches  424;  Conservative    0;  Mismatches 209;  Indels    3;  Gaps    2;

Db       1471 CCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGCTGGTT 1530
              |||||||  || || || ||||||   |  ||||||||  |||| |||| |||| ||||
Qy        703 CCCTTCACCGAGGCCTCCATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTA 762

Db       1531 CACATGATCAACTGGGCGAAGAGGGTGCCAGGCTTTGTGGATTTGACCCTCCATGATCAG 1590
              |||||||||| |||||| |||| | | || |||||||||||  | | |||    || ||
Qy        763 CACATGATCAGCTGGGCCAAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAA 822

Db       1591 GTCCACCTTCTAGAATGTGCCTGGCTAGAGATCCTGATGATTGGTCTCGTCTGGCGCTCC 1650
              || |  ||  | ||  |    ||| | ||| |  | ||||| || ||  | ||||||||
Qy        823 GTGCGGCTCTTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCA 882

Db       1651 ATGGAGCACCCAGTGAAGCTACTGTTTGCTCCTAACTTGCTCTTGGACAGGAACCAGGGA 1710
              || || ||||| |  |||||  | ||||||||  |  |  |  |||||||| |  ||||
Qy        883 ATTGACCACCCCGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGG 942

Db       1711 AAATGTGTAGAGGGCATGGTGGAGATCTTCGACATGCTGCTGGCTACATCATCTCGGTTC 1770
              ||||| ||||| || ||  |||| ||||| |||||||| ||||| ||  | ||  ||||
Qy        943 AAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTT 1002

Db       1771 CGCATGATGAATCTGCAGGGAGAGGAGTTTGTGTGCCTCAAATCTATTATTTTGCTTAAT 1830
              ||   | | || || ||     | || | | | ||  ||||  | || ||  |||| |||
Qy       1003 CGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAAT 1062

Db       1831 TCTGGAGTGTACACATTTCTGTCCAGCACCCTGAAGTCTCTGGAAGAGAAGGACCATATC 1890
              ||  |  ||||| |  |  |  |    |||| | |  ||    |  ||  |||   |  |
Qy       1063 TCCAGTATGTACCCTCTGGTCACAGCGACCCAGGATGCTGAC-AGCAGCCGGAAGCTGGC 1121

Db       1891 CACCGAGTCCTGGACAAGATCACAGACACTTTGATCCACCTGATGGCCAAGGCAGGCCTG 1950
                 |  |  ||| ||    | || ||  ||||| |     |||| ||||||   ||| |
Qy       1122 TCACTTG--CTGAACGCCGTGACCGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATC 1179

Db       1951 ACCCTGCAGCAGCAGCACCAGCGGCTGGCCCAGCTCCTCCTCATCCTCTCCCACATCAGG 2010
               ||   ||||||||   |   ||| ||||| | |||||  |  |||| |||||| |||||
Qy       1180 TCCTCCCAGCAGCAATCCATGCGGCCTGGCTAACCTCCTGATGCTCCTGTCCCACGTCAGG 1239

Db       2011 CACATGAGTAACAAAGGCATGGAGCATCTGTACAGCATGAAGTGCAAGAACGTGGTGCCC 2070
              ||   ||||||||| |||||||| ||||||  || ||||||| |||| || ||||| ||
Qy       1240 CATGCGAGTAACAAGGGCATGGAACATCTGCTCAACATGAAC GCAAAAATGTGCTCCCA 1299

Db       2071 CTCTATGACCTGCTGCTGGAGATGCTGGACGCCCAC 2106
               | |||||||||||||||||||||||| | ||||||
Qy       1300 GTGTATGACCTGCTGCTGGAGATGCTGAATGCCCAC 1335


RESULT     2
ID   US-08-260-452-9 STANDARD; DNA; UNC; 4963 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 9, Application US/08260452.
CC   Sequence 9, Application US/08260452
CC   Patent No. 5650298
CC    GENERAL INFORMATION:
CC       APPLICANT:  Gossen, Manfred
CC       APPLICANT:  Bujard, Hermann
CC       APPLICANT:  Salfeld, Jochen
CC       APPLICANT:  Voss, Jeffrey
CC       TITLE OF INVENTION:  Tight Control of Gene Expression in Eucaryotic
CC       TITLE OF INVENTION:  Cells by Tetracycline-responsive Promoters
CC       NUMBER OF SEQUENCES:  10
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Lahive & Cockfield
CC         STREET:  60 State Street
CC         CITY:  Boston
CC         STATE:  Massachusetts
CC         COUNTRY:  USA
CC         ZIP:  02109-1875
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  ASCII text
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/260,452
CC         FILING DATE:
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  08/076,327
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Giulio A. DeConti, Jr.
CC         REGISTRATION NUMBER:  31,503
CC         REFERENCE/DOCKET NUMBER:  BBI-013CP
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (617) 227-7400
CC         TELEFAX:  (617) 227-5941
CC     INFORMATION FOR SEQ ID NO:  9:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  4963 base pairs
CC         TYPE:  nucleic acid
CC         STRANDEDNESS:  double
CC         TOPOLOGY:  circular
CC       MOLECULE TYPE:  DNA (genomic)
CC       ORIGINAL SOURCE:
CC         ORGANISM:  Human cytomegalovirus
CC       IMMEDIATE SOURCE:
CC         CLONE:  pUHD BGR4
SQ   Sequence 4963 BP; 1259 A; 1300 C; 1278 G; 1126 T; 0 other;

  Query Match             13.7%;  Score 197;  DB 7;  Length 4963;
  Best Local Similarity   66.7%;  Pred. No. 4.16e-135;
  Matches  424;  Conservative    0;  Mismatches 209;  Indels    3;  Gaps    2;

Db       1471 CCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGCTGGTT 1530
              |||||||  || || || ||||||   |  ||||||||  |||| |||| |||| ||||
Qy        703 CCCTTCACCGAGGCCTCCATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTA 762

Db       1531 CACATGATCAACTGGGCGAAGAGGGTGCCAGGCTTTGTGGATTTGACCCTCCATGATCAG 1590
              |||||||||| |||||| |||| | | || |||||||||||  | | |||    || ||
Qy        763 CACATGATCAGCTGGGCCAAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAA 822

Db       1591 GTCCACCTTCTAGAATGTGCCTGGCTAGAGATCCTGATGATTGGTCTCGTCTGGCGCTCC 1650
              || |  ||  | ||  |    ||| | ||| |  | ||||| || ||  | ||||||||
Qy        823 GTGCGGCTCTTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCA 882

Db       1651 ATGGAGCACCCAGTGAAGCTACTGTTTGCTCCTAACTTGCTCTTGGACAGGAACCAGGGA 1710
              || || ||||| |  |||||  | ||||||||  |  |  |  |||||||| |  ||||
Qy        883 ATTGACCACCCCGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGG 942

Db       1711 AAATGTGTAGAGGGCATGGTGGAGATCTTCGACATGCTGCTGGCTACATCATCTCGGTTC 1770
              ||||| ||||| || ||  |||| ||||| |||||||| ||||| ||  | ||  ||||
Qy        943 AAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTT 1002

Db       1771 CGCATGATGAATCTGCAGGGAGAGGAGTTTGTGTGCCTCAAATCTATTATTTTGCTTAAT 1830
              ||   | | || || ||     | || | | | ||  ||||  | || ||  |||| |||
Qy       1003 CGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAAT 1062

Db       1831 TCTGGAGTGTACACATTTCTGTCCAGCACCCTGAAGTCTCTGGAAGAGAAGGACCATATC 1890
              ||  |  ||||| |  |  |  |    |||| | |  ||    |  ||  |||   |  |
Qy       1063 TCCAGTATGTACCCTCTGGTCACAGCGACCCAGGATGCTGAC-AGCAGCCGGAAGCTGGC 1121

Db       1891 CACCGAGTCCTGGACAAGATCACAGACACTTTGATCCACCTGATGGCCAAGGCAGGCCTG 1950
                 |  |  ||| ||    | || ||  ||||| |     |||| ||||||   ||| |
Qy       1122 TCACTTG--CTGAACGCCGTGACCGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATC 1179

Db       1951 ACCCTGCAGCAGCAGCACCAGCGGCTGGCCCAGCTCCTCCTCATCCTCTCCCACATCAGG 2010
               ||   ||||||||   |   ||| ||||| | |||||  |  |||| |||||| |||||
Qy       1180 TCCTCCCAGCAGCAATCCATGCGGCCTGGCTAACCTCCTGATGCTCCTGTCCCACGTCAGG 1239

Db       2011 CACATGAGTAACAAAGGCATGGAGCATCTGTACAGCATGAAGTGCAAGAACGTGGTGCCC 2070
```

```
              |  ||||||| |||||| |||||  || ||||||||||| || ||||| ||
Qy    1240 CATGCGAGTAACAAGGGCATGGAACATCTGCTCAACATGAAGTGCAAAAATGTGCTCCCA 1299

Db    2071 CTCTATGACCTGCTGCTGGAGATGCTGGACGCCCAC 2106
             | ||||||||||||||||||||||||| | ||||||
Qy    1300 GTGTATGACCTGCTGCTGGAGATGCTGAATGCCCAC 1335


RESULT    3
ID    US-08-312-312A-1 STANDARD; DNA; UNC; 3460 BP.
AC    xxxxxx
DT    01-JAN-1900
DE    Sequence 1, Application US/08312312A
CC    Sequence 1, Application US/08312312A
CC    Patent No. 5550024
CC     GENERAL INFORMATION:
CC       APPLICANT:  Rothschild, Max F.
CC       APPLICANT:  Tuggle, Christopher K.
CC       APPLICANT   Jacobson, Carol D.
CC       APPLICANT:  Mileham, Alan J.
CC       APPLICANT:  Plastow, Graham S.
CC       APPLICANT:  Vaske, David A
CC       TITLE OF INVENTION:  Genetic Markers for Pig Litter Size
CC       NUMBER OF SEQUENCES:  9
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Zarley, McKee, Thomte, Voorhees & Sease
CC         STREET:  801 Grand Suite 3200
CC         CITY:  Des Moines
CC         STATE:  Iowa
CC         COUNTRY:  United States
CC         ZIP:  50309
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE.  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.30
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER   US/08/312,312A
CC         FILING DATE:  26-SEP-1994
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 07/961,819
CC         FILING DATE.  19-OCT-1992
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 07/687,708
CC         FILING DATE:  19-APR-1991
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Nebel, Heidi S.
CC         REGISTRATION NUMBER:  37,719
CC         REFERENCE/DOCKET NUMBER:  BRDC-4
CC       TELECOMMUNICATION INFORMATION
CC         TELEPHONE:  515-288 3667
CC         TELEFAX:  515-288 1338
CC     INFORMATION FOR SEQ ID NO:  1:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  3460 base pairs
CC         TYPE:  nucleic acid
CC         STRANDEDNESS:  double
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  DNA (genomic)
CC       HYPOTHETICAL:  NO
CC       ANTI-SENSE:  NO
CC       FEATURE:
CC         NAME/KEY:  misc_feature
CC         LOCATION:  303..304
CC         OTHER INFORMATION   /note= "Gap in DNA sequence of
CC         OTHER INFORMATION:  undetermined length"
CC       FEATURE:
CC         NAME/KEY:  misc_feature
CC         LOCATION:  1197..1198
CC         OTHER INFORMATION:  /note= "Gap in DNA sequence of
CC         OTHER INFORMATION   undetermined length"
CC       FEATURE:
CC         NAME/KEY:  exon
CC         LOCATION:  534..600
SQ   Sequence 3460 BP; 1052 A; 765 C; 634 G; 1009 T; 0 other;

  Query Match              5.0%;  Score 71;  DB 6;  Length 3460;
  Best Local Similarity  83.3%;  Pred. No. 1.13e-34;
  Matches    95;  Conservative    0;  Mismatches  18;  Indels   1;  Gaps   1;

Db     490 TTCATCATGCCCACTTCGTAGCACTTGCGTAGCCGGCAGGC-TGACAGCTCTTCCTCCTG 548
           |||| ||| |||||||||||| ||||| || || ||||||| || |||||||| | ||
Cp     488 TTCACCATTCCCACTTCGTAACACTTCCGAAGTCGGCAGGCCTGGCAGCTCTTGCGCCGG 429

Db     549 TTTTTATTAATTGTTTACTCGTTCGTAGCTGGATACATGTAGTCATTATGTCCT 602
           || ||||| ||||| ||||| || |||||||||| || || |||||||||| |
Cp     428 TTTTTATCGATTGTACACTGATTTGTAGCTGGACAAATATAATCATTATGTCCT 375


RESULT    4
ID    US-08-076-726-15 STANDARD, DNA, UNC, 6244 BP.
AC    xxxxxx
DT    01-JAN-1900
DE    Sequence 15, Application US/08076726.
CC    Sequence 15, Application US/08076726
CC    Patent No. 5464758
CC     GENERAL INFORMATION:
CC       APPLICANT:  Gossen, Manfred
CC       APPLICANT:  Bujard, Hermann
CC       TITLE OF INVENTION.  Tight Control of Gene Expression in
CC       TITLE OF INVENTION:  Eucaryotic Cells by Tetracycline-responsive Promot
  ers
CC       NUMBER OF SEQUENCES:  16
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Sterne, Kessler, Goldstein & Fox
CC         STREET:  1100 New York Avenue, N.W.
CC         CITY:  Washington
CC         STATE:  District of Columbia
CC         COUNTRY:  United States of America
CC         ZIP   20005-3934
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE   PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/076,726
CC         FILING DATE:  14-JUN-1993
CC         CLASSIFICATION:  435
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Esmond, Robert W.
CC         REGISTRATION NUMBER:  32,893
CC         REFERENCE/DOCKET NUMBER:  0942.2490001
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE.  (202)371-2600
CC         TELEFAX:  (202)371-2540
CC         TELEX:  248636 SSK
CC     INFORMATION FOR SEQ ID NO:  15:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH.  6244 base pairs
CC         TYPE:  nucleic acid
CC         STRANDEDNESS:  both
CC         TOPOLOGY:  both
SQ   Sequence 6244 BP; 1483 A; 1749 C; 1610 G; 1402 T; 0 other;

  Query Match              4.4%;  Score 63;  DB 5;  Length 6244;
  Best Local Similarity  66.3%;  Pred. No. 1.10e-28;
  Matches   128;  Conservative    0;  Mismatches  65;  Indels   0;  Gaps   0;

Db    2365 TCTGTGGGGATGAAGTATCAGCTTGTCATTATGGTGTCCTCACCTGTGGGAGCTGTAAGG 2424
           ||||   | ||| | ||||| || | ||| ||||| |||     | ||||    |||||||
Qy     293 TCTGTAGCGATTACGTATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATGTAAGG 352

Db    2425 TCTTCTTTAAAAGGGCAATGGAAGGGTAGTATAACTATTTATGTGTTGGAAGAAATGACT 2484
            ||| ||||||||    ||  |||| ||  || | ||| | ||| | |   | |||| | |
Qy     353 CCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCAGT 412

Db    2485 GCATTGTTGATAAAATCCGCAGGAAAAACTGCCGGCCGTGTGGCCTTAGAAAGTGCTCTC 2544
           | |    | ||||||| |||    | || | ||||| ||| || || ||| | ||||| |
Qy     413 GTACAATCGATAAAAACCGGGCAAGAGCTGGCAGGCCTCCCGACTTCGGAAGTCGTTACG 472

Db    2545 AAGCTGGCATGGT 2557
           |||   || |||||
Qy     473 AAGTGGGAATGGT 485

RESULT     5
ID   US-08-260-452-8 STANDARD; DNA; UNC; 6244 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 8, Application US/08260452
CC   Sequence 8, Application US/08260452
CC   Patent No. 5650298
CC    GENERAL INFORMATION:
CC       APPLICANT:  Gossen, Manfred
CC       APPLICANT:  Bujard, Hermann
CC       APPLICANT:  Salfeld, Jochen
CC       APPLICANT:  Voss, Jeffrey
CC       TITLE OF INVENTION:  Tight Control of Gene Expression in Eucaryotic
CC       TITLE OF INVENTION:  Cells by Tetracycline-responsive Promoters
CC       NUMBER OF SEQUENCES:  10
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Lahive & Cockfield
CC         STREET:  60 State Street
CC         CITY:  Boston
CC         STATE:  Massachusetts
CC         COUNTRY:  USA
CC         ZIP:  02109-1875
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  ASCII text
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/260,452
CC         FILING DATE:
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  08/076,327
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Giulio A. DeConti, Jr
CC         REGISTRATION NUMBER:  31,503
CC         REFERENCE/DOCKET NUMBER:  BBI-013CP
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (617) 227-7400
CC         TELEFAX:  (617) 227-5941
CC    INFORMATION FOR SEQ ID NO:  8:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  6244 base pairs
CC         TYPE:  nucleic acid
CC         STRANDEDNESS:  double
CC         TOPOLOGY:  circular
CC       MOLECULE TYPE:  DNA (genomic)
CC       ORIGINAL SOURCE:
CC         ORGANISM:  Human cytomegalovirus
CC         STRAIN:  Towne (hCMV)
CC       IMMEDIATE SOURCE:
CC         CLONE:  pUHD BGR3
SQ   Sequence 6244 BP, 1483 A, 1749 C, 1610 G, 1402 T, 0 other;

  Query Match              4.4%,  Score 63,  DB 7,  Length 6244;
  Best Local Similarity  66.3%;  Pred. No. 1.10e-28;
  Matches  123;  Conservative    0;  Mismatches  65;  Indels    0;  Gaps    0;

Db    2365 TCTGTGGGCATGAAGCATCAGCCTGTCATTATGGTGTCCTCACCTGTGGGAGCTGTAAGG 2424
           ||||   | ||| | ||||| || | ||| ||||| |||      | ||||    | |||||||
Qy     293 TCTCTACCGATTACGGCATCGGGATATCACTATGGAGTCTGCCCTGTGAAGGATGTAAGG 352
```

```
Db    2425 TCTTCTTTAAAAGGGAATGCAAGGGCAGCATAACTATTTATGTGCCTGGAAGAAATGACT 2484
            ||| ||||||||     ||   |||| ||   || | ||| | ||| | |   | |||| | |
Qy     353 CCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCAGT 412

Db    2485 GCATTGTTGATAAAATCCGCAGGAAAAACTGCCGGCCGTGTGGCCTTAGAAAGTGCTCTC 2544
           | |    | ||||||| |||    | || | ||||| ||| || || ||| | ||||| |
Qy     413 GTACAATCGATAAAAACCGGGCAAGAGCTGGCAGGCCTCCCGACTTCGGAAGTCGTTACG 472

Db    2545 AAGCTGGCATGGT 2557
           |||   || |||||
Qy     473 AAGTGGGAATGGT 485

RESULT     6
ID   US-08-232-463-14 STANDARD; DNA; UNC; 7218 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 14, Application US/08232463.
CC   Sequence 14, Application US/08232463
CC   Patent No. 5670367
CC    GENERAL INFORMATION:
CC       APPLICANT:  DORNER, F.
CC       APPLICANT:  SCHEIFLINGER, F
CC       APPLICANT:  FALKNER, F. G.
CC       TITLE OF INVENTION:  RECOMBINANT FOWLPOX VIRUS
CC       NUMBER OF SEQUENCES:  52
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Foley & Lardner
CC         STREET:  1800 Diagonal Road, Suite 500
CC         CITY:  Alexandria
CC         STATE:  VA
CC         COUNTRY:  USA
CC         ZIP:  22313-0299
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/232,463
CC         FILING DATE:
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/935,313
CC         FILING DATE:
CC         APPLICATION NUMBER:  EP 91 114 300.6
CC         FILING DATE:  26-AUG-1991
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  BENT, Stephen A.
CC         REGISTRATION NUMBER:  29,768
CC         REFERENCE/DOCKET NUMBER:  30472/114 IMMU
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (703)836-9300
CC         TELEFAX:  (703)683-4109
CC         TELEX:  899149
CC    INFORMATION FOR SEQ ID NO:  14:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  7218 base pairs
CC         TYPE:  nucleic acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       IMMEDIATE SOURCE:
CC         CLONE:  pTZgpt-F1s
SQ   Sequence 7218 BP, 1944 A, 1491 C, 1486 G, 1929 T, 368 other;

  Query Match              4.3%,  Score 62;  DB 7,  Length 7218;
  Best Local Similarity   1.4%;  Pred. No. 6.06e-28;
  Matches    5;  Conservative  205;  Mismatches 148;  Indels    0;  Gaps    0;

Db    1083 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1142
           ::::: : ::.      .  .  .   .: :.  ........   .... .:   :  .   : :   ::
Qy     612 CTTTGTATTGGGGCGCTGTGGTCTGCACTTTTCTTGGCCTTGCGCCCTACACCCTACCTC 553
```

```
Db   1143 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1202
           ::: ::  : :::::::  ..... . .      .  .  ... . ..........     ..
Cp    552 CTCGTCGGCACTTCTTGTCTCGGCACAAGGCGGTACCCACATCTCTTCTCCGGGAGCC 493

Db   1203 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1262
            : :::: :: ::::: :::: :   : :::::      ::  :    .::  :  :::::  .
Cp    492 ACACTTCACCATTCCCACTTCGTAACACTTCCGAAGTCGGCAGCCTGGCAGCTCTTCCC 433

Db   1263 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1322
           ::  ::::: ..  .. . . ..  ... .  .:    :    : :  :: :: : :::::
Cp    432 CCCGTTTTTATCGATTCTACACTGATTTGTAGCTGGACAAATATAATCATTATCTCCTTC 373

Db   1323 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1382
             : :::::::         :... . ...... . .  ..  .... .  .  : ...
Cp    372 AATGCTTCTTTTAAAAAAGGCCTTACATCCTTCACACGACCAGACTCCATAGTGATAICC 313

Db   1383 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYGTACCA 1440
           :  :  :  :  :: :: :    :   : :      :    : :::::::::    :::| ||||
Cp    312 CGATGCGTAATCGCTGCAGACAGCGCAGAAGTGAGCATCCCTCTTTGAACCTCGACCA 255


RESULT    7
ID   5223606-1 STANDARD; DNA; UNC; 3238 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Patent No. 5223606.
CC  Patent No. 5223606
CC      APPLICANT   BLAUDIN DE THE, HUGHES;MARCHIO, AGNES;TIOLLAIS,
CC  PIERRE;DEJEAN, ANNE
CC      TITLE OF INVENTION  STEROID/THYROID HORMONE RECEPTOR-RELATED
CC  PROTEIN INAPPROPRIATELY EXPRESSED IN HUMAN HEPATOCELLULAR CARCINOMA
CC      NUMBER OF SEQUENCES: 11
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  US/07/134 120
CC        FILING DATE: 17-DEC-1987
CC      PRIOR APPLICATION DATA:
CC  SEQ ID NO:1:
CC        LENGTH: 2989
SQ   Sequence 3238 BP; 900 A; 648 C; 600 G; 841 T; 249 other;

  Query Match             3.8%;  Score 55;  DB 1;  Length 2989;
  Best Local Similarity  65.0%;  Pred. No. 9.23e-23.
  Matches   130;  Conservative    0;  Mismatches  69;  Indels   1;  Gaps   1;

Db    560 CTGCTTCGTCTGCCAGGACAAATCATCAGGGTACCACTATGGGGTCAGCGCCTGTGAGGG 619
          ||||   ||||||   ||   |   |||| || || |||||||| ||| |   | ||||| ||
Qy    285 CTGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGG 344

Db    620 ATGTAAGGGCTTTTTCGGCAGAAGTATTCAGAAGAATATGATTTACACTTGTCACCGAGA 679
          ||||||||| ||||||      ||||| |||||        |||      ||| | |||||
Qy    345 ATGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTAC 404

Db    680 TAAGAACTGTGTTATTAATAAAAGTCACCAGGAATCGATGCCAATACTGTCGACTCCAGA 739
           ||  | |||   ||  ||||||   |   |  |  ||  | |||||   ||| ||||| | ||
Qy    405 AAATCAGTGTACAATCGATAAAAACCGGC-GCAAGAGCTGCCAGGCCTGCCGACTTCGGA 463

Db    740 AGTGCTTTGAAGTGGGAATG 759
          |||| |  |||||||||||
Qy    464 AGTGTTACGAAGTGGGAATG 483


RESULT    8
ID   US-08-629-939-1 STANDARD; DNA; UNC; 3014 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 1, Application US/08629939
CC   Sequence 1, Application US/08629939
CC   Patent No. 5645995
CC    GENERAL INFORMATION:
CC      APPLICANT:  Kioback  Dirk G
CC      TITLE OF INVENTION:  METHODS FOR DIAGNOSING AN INCREASED
CC      TITLE OF INVENTION:  RISK OF BREAST OR OVARIAN CANCER
CC      NUMBER OF SEQUENCES:  14
```

```
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  SUGHRUE, MION, ZINN, MACPEAK & SEAS
CC        STREET:  2100 Pennsylvania Avenue, N.W., Suite 800
CC        CITY:  Washington, D.C.
CC        STATE:  D.C.
CC        COUNTRY:  U.S.A.
CC        ZIP:  20037
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  US/08/629,939
CC        FILING DATE:  12-APRIL-1996
CC        CLASSIFICATION:  435
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  KIT, Gordon
CC        REGISTRATION NUMBER:  30,764
CC        REFERENCE/DOCKET NUMBER:  A-6612
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  (202) 293-7060
CC        TELEFAX:  (202) 293-7860
CC    INFORMATION FOR SEQ ID NO:  1:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  3014 base pairs
CC        TYPE:  nucleic acid
CC        STRANDEDNESS:  single
CC        TOPOLOGY:  linear
CC      MOLECULE TYPE:  cDNA
CC      HYPOTHETICAL:  NO
SQ   Sequence 3014 BP; 623 A; 939 C; 827 G; 625 T; 0 other;

  Query Match             3.7%;  Score 53;  DB 7;  Length 3014;
  Best Local Similarity  63.7%;  Pred. No. 2.30e-21;
  Matches   123;  Conservative    0;  Mismatches  70;  Indels   0;  Gaps   0;

Db   1881 TCTGTGGGGATGAAGCATCAGGCTGTCATTATGGTGTCCTTACCTGTGGGAGCTGTAAGG 1940
          ||||  | ||| | ||||| || | ||| ||||| |||     | ||||    | ||||||
Qy    293 TCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATGTAAGG 352

Db   1941 TCTTCTTTAAGAGGGCAATGGAAGGGCAGCACAACTACTTATGTGCTGGAAGAAATCACT 2000
           ||| ||||| ||     ||  |||| ||  |   | ||  | |||  |  |   | |||| | |
Qy    353 CCTTTTTTAAAAGAACCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCAGT 412

Db   2001 GCATCGTTGATAAAATCCGCAGAAAAAACTGCCCAGCATGTCGCCTTAGAAAGTCCTGTC 2060
          | |    | ||||||| |||   | |  | ||||  || || || ||| | |||||  |
Qy    413 GTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTACG 472

Db   2061 AGGCTGGCATGGT 2073
          | |   || |||||
Qy    473 AAGTGGGAATGGT 485


RESULT    9
ID   5260432-1 STANDARD; DNA; UNC; 1707 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Patent No. 5260432.
CC  Patent No. 5260432
CC      APPLICANT: TAKAKU, FUMIMARO;ISHIKAWA, TAKASHI;IMAWARI, MICHIO;
CC  EVANS, RONALD M.;UMESONO, KAZUHIKO
CC      TITLE OF INVENTION: HUMAN GAMMA RETINOIC ACID RECEPTOR DNA
CC      NUMBER OF SEQUENCES: 4
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  US/07/370,407
CC        FILING DATE: 22-JUN-1989
CC  SEQ ID NO:1:
CC        LENGTH: 1576
SQ   Sequence 1707 BP; 346 A; 478 C; 467 G; 285 T; 131 other;

  Query Match             3.6%;  Score 51;  DB 1;  Length 1576;
  Best Local Similarity  65.8%;  Pred. No. 6.30e-20;
  Matches   127,  Conservative    0,  Mismatches  64,  Indels   2,  Gaps   2,

Db    473 GTGTGCAATGACAAGTCCTCTGGCTACCACTATGGGGTCAGCTCTTGTGAAGGCTGCAAG 532
          || ||||  ||  |  | || || || |||||||| ||| | || |||||||| || |||
Qy    292 GTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATGTAAG 351

Db    533 GGCTTCTTTCGGCGAAGCATCCAGAAGAACATGGTGTACACGTGTCACCGCGACAAAA-A 591
          | ||| |||    ||||||| ||     | |  |  || |  |||| | || |||||  |
Qy    352 GCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTC-CAGCTACAAATCA 410

Db    592 CTGTATCATCAACAAGGTGACCAGGAATCGCTGCCAGTACTGCCGGCTACAGAAGTGCTT 651
           ||||  ||| | ||        | ||  ||||||||  |||||| || | |||||| |
Qy    411 GTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTA 470

Db    652 CGAAGTGGGCATG 664
          ||||||||| |||
Qy    471 CGAAGTGGGAATG 483


RESULT   10
ID   US-08-087-151-1 STANDARD; DNA; UNC; 387 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 1, Application US/08087151
CC   Sequence 1, Application US/08087151
CC   Patent No. 5468624
CC    GENERAL INFORMATION:
CC      APPLICANT:  E. Brad Thompson
CC      APPLICANT:  Lynne V. Nazareth
CC      TITLE OF INVENTION:  ANTI-TUMOR ACTIVITY OF A MODIFIED FRAGMENT
CC      TITLE OF INVENTION:  OF THE GLUCOCORTICOID RECEPTOR
CC      NUMBER OF SEQUENCES:  3
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  Arnold, White & Durkee
CC        STREET:  P.O. Box 4433
CC        CITY:  Houston
CC        STATE:  Texas
CC        COUNTRY:  USA
CC        ZIP:  77210
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  WordPerfect 5.1
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  US/08/087,151
CC        FILING DATE:  July 7, 1993
CC        CLASSIFICATION:  435
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  Hodgins, Daniel S.
CC        REGISTRATION NUMBER:  31,026
CC        REFERENCE/DOCKET NUMBER:  UTSG:141/HOD
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  512-320-7200
CC        TELEFAX:  512-474-7577
CC        TELEX:  No. 5468624 Applicable
CC    INFORMATION FOR SEQ ID NO:  1:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  387 base pairs
CC        TYPE:  nucleic acid
CC        STRANDEDNESS:  double
CC        TOPOLOGY:  linear
CC      MOLECULE TYPE:  genomic DNA
SQ   Sequence 387 BP; 109 A; 92 C; 88 G; 98 T; 0 other;

  Query Match             3.3%;  Score 47;  DB 5;  Length 387;
  Best Local Similarity  63.0%;  Pred. No. 4.37e-17;
  Matches   126,  Conservative    0,  Mismatches  73   Indels   1;  Gaps   1;

Db    188 TCTGCCTGGTGTGCTCTGATGAAGCTTCAGGATGTCATTATGGAGTCTTAACTTGTGGAA 247
          |||||   || |||   ||| | || || |||| ||| ||||||||||   | |||| |
Qy    284 TCTGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAG 343
```

```
Db    248 GCTGTAAAGTTTTCTTCAAAAGAGCAGTGGAAGGACAGCACAATTACCTATGTGCTGGAA 307
          | ||||| |  || || ||||||    |  |||||||  |  ||||  | ||| | |  |
Qy    344 GATGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTA 403

Db    308 GGAATGATTGCATCATCGCGATAAAATT-CGAAAAAACTGCCCAGCATGCCGCTATCGAA 366
            ||| | || |  ||||  | |||    || || | |||||  || |||||   ||| |
Qy    404 CAAATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGA 463

Db    367 AATGTCTTCAGGCTGGAATG 386
          | |||    | |  ||||||
Qy    464 AGTGTTACGAAGTGGGAATG 483


RESULT   11
ID   US-08-561-072-1 STANDARD; DNA; UNC; 387 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 1, Application US/08561072.
CC   Sequence 1, Application US/08561072
CC   Patent No. 5571791
CC    GENERAL INFORMATION:
CC      APPLICANT:  E. Brad Thompson
CC      APPLICANT:  Lynne V. Nazareth
CC      TITLE OF INVENTION:  ANTI-TUMOR ACTIVITY OF A MODIFIED FRAGMENT
CC      TITLE OF INVENTION:  OF THE GLUCOCORTICOID RECEPTOR
CC      NUMBER OF SEQUENCES:  3
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  Arnold, White & Durkee
CC        STREET:  P.O. Box 4433
CC        CITY:  Houston
CC        STATE:  Texas
CC        COUNTRY:  USA
CC        ZIP:  77210
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  WordPerfect 5.1
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  US/08/561,072
CC        FILING DATE:  20-NOV-1995
CC        CLASSIFICATION:  530
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  08/087,151
CC        FILING DATE:  July 7, 1993
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  Hodgins, Daniel S.
CC        REGISTRATION NUMBER:  31,026
CC        REFERENCE/DOCKET NUMBER:  UTSG:141/HOD
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  512-320-7200
CC        TELEFAX:  512-474-7577
CC        TELEX:  No. 5571791 Applicable
CC    INFORMATION FOR SEQ ID NO:  1:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  387 base pairs
CC        TYPE:  nucleic acid
CC        STRANDEDNESS:  double
CC        TOPOLOGY:  linear
CC      MOLECULE TYPE:  genomic DNA
SQ   Sequence 387 BP; 109 A; 92 C; 88 G; 98 T; 0 other;

  Query Match             3.3%;  Score 47;  DB 6;  Length 387;
  Best Local Similarity  63.0%;  Pred. No. 4.37e-17;
  Matches   126;  Conservative    0;  Mismatches  73;  Indels   1;  Gaps   1;

Db    188 TCTGCCTGGTGTGCTCTGATGAAGCTTCAGGATGTCATTATGGAGTCTTAACTTGTGGAA 247
          |||||   || |||   ||| | || || |||| ||| ||||||||||   | |||| |
Qy    284 TCTGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAG 343

Db    248 GCTGTAAAGTTTTCTTCAAAAGAGCAGTGGAAGGACAGCACAATTACCTATGTGCTGGAA 307
          | ||||| |  || || ||||||    |  |||||||  |  ||||  | ||| | |  |
Qy    344 GATGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTA 403
```

```
Db    308 GGAATGATTGCATCATCGGGATAAAATT-CGAAAAAACTGCCCAGCATGCCGCTATCGAA 366
            ||| | || |  ||||  | |||    || || |  ||||  || |||||   ||| |
Qy    404 CAAATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGA 463

Db    367 AATGTCTTCAGGCTGGAATG 386
          | |||    | |  ||||||
Qy    464 AGTGTTACGAAGTGGGAATG 483


RESULT   12
ID   PCT-US92-02320A-3 STANDARD, DNA, UNC, 2928 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 3, Application PC/TUS9202320A.
CC   Sequence 3, Application PC/TUS9202320A
CC    GENERAL INFORMATION:
CC      APPLICANT:  Sloan-Kettering Institute, For Cancer Research
CC      TITLE OF INVENTION:  METHODS FOR DETECTION AND TREATMENT OF CANCER
CC      NUMBER OF SEQUENCES:  6
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  COOPER & DUNHAM
CC        STREET:  30 ROCKEFELLER PLAZA
CC        CITY:  NEW YORK
CC        STATE:  NEW YORK
CC        COUNTRY:  U.S.A.
CC        ZIP:  10112
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  PatentIn Release #1.24
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  PCT/US92/02320A
CC        FILING DATE:
CC        CLASSIFICATION:
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US 675,084
CC        FILING DATE:  22-MAR-1991
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US 673,838
CC        FILING DATE:  22-MAR-1991
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  WHITE, JOHN P
CC        REGISTRATION NUMBER:  28,678
CC        REFERENCE/DOCKET NUMBER:  38694-PCT
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  (212) 977-9550
CC        TELEFAX:  (212) 644-0525
CC        TELEX:  (212) 422523 COOP UI
CC    INFORMATION FOR SEQ ID NO:  3:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  2928 base pairs
CC        TYPE:  nucleic acid
CC        STRANDEDNESS:  single
CC        TOPOLOGY:  linear
CC      MOLECULE TYPE:  cDNA
CC      IMMEDIATE SOURCE:
CC        CLONE:  hRAR ALPHA
CC      FEATURE:
CC        NAME/KEY:  CDS
CC        LOCATION:  103..1488
CC        OTHER INFORMATION:
SQ   Sequence 2928 BP; 592 A; 1011 C; 783 G; 546 T; 0 other;

  Query Match             2.9%;  Score 42;  DB 10;  Length 2928;
  Best Local Similarity  60.6%;  Pred. No. 1.32e-13;
  Matches  120;  Conservative    0;  Mismatches  78;  Indels   0;  Gaps    0;

Db    364 TGCTTTGTCTGTCAGGACAAGTCCTCAGGCTACCACTATGGGGTCAGCGCCTGTGAGGGC 423
          |||  ||||||    ||  |  | || || || |||||||| ||| |  | ||||| ||
Qy    286 TGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGA 345

Db    424 TGCAAGGGCTTCTTCCGCCGCAGCATCCAGAAGAACATGGTGTACACGTGTCACCGGGAC 483
          || |||| ||| ||     | ||||| ||     | |  |  || |  ||||
Qy    346 TGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACA 405

Db    484 AAGAACTGCATCATCAACAAGGTGACCCGGAACCGCTGCCAGTACTGCCGACTGCAGAAG 543
          ||  | || |  ||| | ||       || ||  ||||||||  ||||||||| | ||||
Qy    406 AATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAG 465

Db    544 TGCTTTGAAGTGGGCATG 561
          || |  |||||||| |||
Qy    466 TGTTACGAAGTGGGAATG 483


RESULT   13
ID   5171671-1 STANDARD; DNA; UNC; 3185 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Patent No. 5171671.
CC  Patent No. 5171671
CC       APPLICANT: EVANS, RONALD M.;ONG, ESTELITA S.;SEGUI,
CC  PRUDIMAR S.;THOMPSON, CATHERINE C.;UEMSONO, KAZUHIKO
CC  GUGUERE, VINCENT
CC       TITLE OF INVENTION: RETINOIC ACID RECEPTOR COMPOSITION
CC       NUMBER OF SEQUENCES: 2
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER: US/07/546,256
CC         FILING DATE: 06-AUG-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER: 276,536
CC         FILING DATE: 30-NOV-1988
CC         APPLICATION NUMBER: 128,331
CC         FILING DATE: 02-DEC-1987
CC  SEQ ID NO:1:
CC         LENGTH: 2940
SQ   Sequence 3185 BP; 594 A; 1010 C; 790 G; 546 T; 245 other;

  Query Match             2.9%;  Score 42;  DB 1;  Length 2940;
  Best Local Similarity  60.6%;  Pred. No. 1.32e-13;
  Matches  120;  Conservative    0;  Mismatches  78;  Indels   0;  Gaps    0;

Db    364 TGCTTTGTCTGTCAGGACAAGTCCTCAGGCTACCACTATGGGGTCAGCGCCTGTGAGGGC 423
          |||  ||||||    ||  |  | || || || |||||||| ||| |  | ||||| ||
Qy    286 TGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGA 345

Db    424 TGCAAGGGCTTCTTCCGCCGCAGCATCCAGAAGAACATGGTGTACACGTGTCACCGGGAC 483
          || |||| ||| ||     | ||||| ||     | |  |  || |  ||||
Qy    346 TGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACA 405

Db    484 AAGAACTGCATCATCAACAAGGTGACCCGGAACCGCTGCCAGTACTGCCGACTGCAGAAG 543
          ||  | || |  ||| | ||       || ||  ||||||||  ||||||||| | ||||
Qy    406 AATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAG 465

Db    544 TGCTTTGAAGTGGGCATG 561
          || |  |||||||| |||
Qy    466 TGTTACGAAGTGGGAATG 483


RESULT   14
ID   PCT-US92-02320A-1 STANDARD; DNA; UNC; 3036 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 1, Application PC/TUS9202320A.
CC   Sequence 1, Application PC/TUS9202320A
CC    GENERAL INFORMATION:
CC      APPLICANT:  Sloan-Kettering Institute, For Cancer Research
CC      TITLE OF INVENTION:  METHODS FOR DETECTION AND TREATMENT OF CANCER
CC      NUMBER OF SEQUENCES:  6
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  COOPER & DUNHAM
CC        STREET:  30 ROCKEFELLER PLAZA
CC        CITY:  NEW YORK
CC        STATE:  NEW YORK
CC        COUNTRY:  U.S.A.
CC        ZIP:  10112
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  PatentIn Release #1.24
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  PCT/US92/02320A
CC        FILING DATE:
CC        CLASSIFICATION:
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US 675,084
CC        FILING DATE:  22-MAR-1991
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US 673,838
CC        FILING DATE:  22-MAR-1991
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  WHITE, JOHN P
CC        REGISTRATION NUMBER:  28,678
CC        REFERENCE/DOCKET NUMBER:  38694-PCT
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  (212) 977-9550
CC        TELEFAX:  (212) 644-6525
CC        TELEX:  (212) 422523 COOP UI
CC    INFORMATION FOR SEQ ID NO:  1:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  3036 base pairs
CC        TYPE:  nucleic acid
CC        STRANDEDNESS:  single
CC        TOPOLOGY:  linear
CC      MOLECULE TYPE:  cDNA
CC      IMMEDIATE SOURCE:
CC        CLONE:  MYL-RAR
CC      FEATURE:
CC        NAME/KEY:  CDS
CC        LOCATION:  67..2457
CC        OTHER INFORMATION:
SQ   Sequence 3036 BP; 605 A; 1018 C; 890 G; 523 T; 0 other;

  Query Match              2.9%;  Score 42;  DB 10;  Length 3036;
  Best Local Similarity   60.6%;  Pred. No. 1.32e-13;
  Matches   120;  Conservative    0;  Mismatches   78;  Indels    0;  Gaps    0;

Db   1333 TGCTTTGTCTGTCAGGACAAGTCCTCAGGCTACCACTATGGGGTCAGCGCCTGTGAGGGC 1392
          |||  ||||||    ||  |  | || || || |||||||| ||| |  | ||||| ||
Qy    286 TGCGGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGA 345

Db   1393 TGCAAGGGCTTCTTCCGCCGCAGCATCCAGAAGAACATGGTGTACACGTGTCACCGGGAC 1452
          || |||| ||| ||     |  |||| ||      | |  | || |  ||||
Qy    346 TGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACA 405

Db   1453 AAGAACTGCATCATCAACAAGGTGACCCGGAACCGCTGCCAGTACTGCCGACTGCAGAAG 1512
          ||  | || |  ||| | ||       || ||  ||||||||  ||||||||| | ||||
Qy    406 AATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAG 465

Db   1513 TGCTTTGAAGTGGGCATG 1530
          || |  |||||||| |||
Qy    466 TGTTACGAAGTGGGAATG 483

RESULT    15
ID   PCT-US91-00399-5 STANDARD; DNA; UNC; 1662 BP
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 5, Application PC/TUS9100399.
CC   Sequence 5, Application PC/TUS9100399
CC    GENERAL INFORMATION:
CC      APPLICANT:  Mangelsdorf, Dr., David J.
CC      APPLICANT:  Evans Dr., Ronald M.
CC      TITLE OF INVENTION:  RETINOID RECEPTOR COMPOSITIONS AND
CC      TITLE OF INVENTION:  METHODS
CC      NUMBER OF SEQUENCES:  7
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  FITCH, EVEN, TABIN & FLANNERY
CC        STREET:  135 South LaSalle Street, Suite 900
CC        CITY:  Chicago
CC        STATE:  Illinois
CC        COUNTRY:  USA
CC        ZIP:  60603
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  PCT/US91/00399
CC        FILING DATE:  19910122
CC        CLASSIFICATION:  435
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  07/478,071
CC        FILING DATE:  09-FEB-1990
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  Reiter, Mr., Stephen E.
CC        REGISTRATION NUMBER:  31192
CC        REFERENCE/DOCKET NUMBER:  50852
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  (619) 552-1311
CC        TELEFAX:  (619) 552-0095
CC        TELEX:  20 6566 PATLAW CGO
CC    INFORMATION FOR SEQ ID NO:  5:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  1662 base pairs
CC        TYPE:  NUCLEIC ACID
CC        STRANDEDNESS:  single
CC        TOPOLOGY:  linear
CC      MOLECULE TYPE:  DNA (genomic)
CC      IMMEDIATE SOURCE:
CC        CLONE:  RXR MOUSE GAMMA
CC      FEATURE:
CC        NAME/KEY:  CDS
CC        LOCATION:  join(123..1514)
SQ   Sequence 1662 BP; 397 A; 483 C; 441 G; 341 T; 0 other;

  Query Match              2.6%;  Score 38;  DB 9;  Length 1662;
  Best Local Similarity   60.1%;  Pred. No. 6.79e-11;
  Matches   113;  Conservative    0;  Mismatches   75;  Indels    0;  Gaps    0;

Db    531 CACATCTGTGCCATCTGTGGGACAGATCCTCAGGGAAGCACTACGGTGTGTACAGCTGT 590
          ||| |||| ||  ||||  | ||     | || ||  | ||||| || || |     |||
Qy    280 CACTTCTGCGGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGT 339

Db    591 GAAGGTTGCAAAGGCTTCTTCAAAAGGACCATCAGGAAAGATCTCATCTACACCTGTCGG 650
          ||||| || || | ||| || ||||| | |||      | ||      || |  ||||
Qy    340 GAAGGATGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCA 399

Db    651 GATAACAAAGATTGTCTCATCGACAAGCGCCAGCGCAACCGCTGCCAGTACTGTCGCTAC 710
          | ||  ||  | |||   ||||| ||   || ||||||  ||||||||  ||| ||
Qy    400 GCTACAAATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTT 459

Db    711 CAGAAGTG 718
          | ||||||
Qy    460 CGGAAGTG 467

Search completed. Thu Dec 18 18.22.52 1997
Job time : 74 secs.
```

```
**********************************************************************

                         MPSRCH (TM)

**********************************************************************

         Release 2.1D John F. Collins, Biocomputing Research Unit
         Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
                Distribution rights by IntelliGenetics, Inc.

MPsrch_pp   protein - protein database search, using Smith-Waterman algorithm

Run on:            Thu Dec 18 09:06:46 1997; MasPar time 11.09 Seconds
                                             799.225 Million cell updates/sec
Tabular output not generated.

Title:             >US-08-826-361A-21
Description:       (1-418) from US08826361A.pep
Perfect Score:     3116
Sequence           1 MNYSIPSNVTNLEGGPGRQT.........SMRLANLLMLLSHVRHARSA 418

Scoring table:     PAM 150
                   Gap 11

Searched           59021 seqs, 21210388 residues

Post-processing:   Minimum Match 0%
                   Listing first 45 summaries

Database:          swiss-prot34
                     1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
                     8:part8 9:part9 10:part10 11:part11

Statistics:        Mean 49.391; Variance 92.280; scale 0.600

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution

                              SUMMARIES

                 %
Result         Query
  No   Score   Match Length DB  ID          Description            Pred. No.
-----------------------------------------------------------------------------
    1   1473   47.3    595   3  ESTR_HUMAN  ESTROGEN RECEPTOR (ER  0.00e+00
    2   1472   47.2    589   3  ESTR_CHICK  ESTROGEN RECEPTOR (ER  0.00e+00
    3   1455   46.7    599   3  ESTR_MOUSE  ESTROGEN RECEPTOR (ER  0.00e+00
    4   1448   46.5    600   3  ESTR_RAT    ESTROGEN RECEPTOR (ER  0.00e+00
    5   1381   44.3    575   3  ESTR_ORYLA  ESTROGEN RECEPTOR (ER  8.70e-303
    6   1327   42.6    574   3  ESTR_SALIR  ESTROGEN RECEPTOR (ER  3.72e-299
    7   1314   42.2    585   3  ESTR_SALSA  ESTROGEN RECEPTOR (ER  6.15e-296
    8    879   28.2    314   3  ESTR_OREAU  ESTROGEN RECEPTOR (ER  1.58e-177
    9    736   23.6    433   3  ERR2_HUMAN  STEROID HORMONE RECEP  1.84e-142
   10    725   23.3    521   3  ERR1_HUMAN  STEROID HORMONE RECEP  3.05e-140
   11    681   21.9    923   7  PRGR_MOUSE  PROGESTERONE RECEPTOR  4.47e-129
   12    666   21.4    933   7  PRGR_HUMAN  PROGESTERONE RECEPTOR  1.94e-125
   13    662   21.2    930   7  PRGR_RABIT  PROGESTERONE RECEPTOR  1.73e-124
   14    614   19.7    777   4  GCR_HUMAN   GLUCOCORTICOID RECEPT  6.80e-113
   15    608   19.5    776   4  GCR_XENLA   GLUCOCORTICOID RECEPT  1.89e-111
   16    597   19.2    786   7  PRGR_CHICK  PROGESTERONE RECEPTOR  8.24e-109
   17    586   18.8    783   4  GCR_MOUSE   GLUCOCORTICOID RECEPT  3.57e-106
   18    579   18.6    133   3  ESTR_BOVIN  ESTROGEN RECEPTOR (ER  1.69e-104
   19    577   18.5    771   4  GCR_CAVPO   GLUCOCORTICOID RECEPT  5.10e-104
   20    577   18.5    795   4  GCR_RAT     GLUCOCORTICOID RECEPT  5.10e-104
   21    562   18.0    981   6  MCR_RAT     MINERALOCORTICOID REC  1.35e-100
   22    557   17.9    899   1  ANDR_MOUSE  ANDROGEN RECEPTOR.     3.05e-99
   23    554   17.8    709   1  ANDR_RABIT  ANDROGEN RECEPTOR (FR  1.58e-98
   24    556   17.8    902   1  ANDR_RAT    ANDROGEN RECEPTOR.     5.27e-99
   25    553   17.7    919   1  ANDR_HUMAN  ANDROGEN RECEPTOR.     2.74e-98
   26    544   17.5    758   4  GCR_ONCMY   GLUCOCORTICOID RECEPT  4.80e-96
   27    527   16.9    984   6  MCR_HUMAN   MINERALOCORTICOID REC  3.69e-91
   28    517   16.6    111   3  ESTR_SHEEP  ESTROGEN RECEPTOR (ER  9.72e-90
   29    505   16.2    458   8  RRXB_RAT    RETINOIC ACID RECEPTO  6.68e-87
   30    505   16.2    520   8  RRXB_MOUSE  RETINOIC ACID RECEPTO  6.68e-87
   31    506   16.2    533   8  RRXB_HUMAN  RETINOIC ACID RECEPTO  3.88e-87
   32    501   16.1    597   6  NGFI_RAT    NERVE GROWTH FACTOR I  5.88e-86
   33    499   16.0    598   6  NGFI_CANFA  ORPHAN NUCLEAR RECEPT  1.74e-85
   34    494   15.9    467   8  RRXA_RAT    RETINOIC ACID RECEPTO  2.63e-84
   35    494   15.9    467   8  RRXA_MOUSE  RETINOIC ACID RECEPTO  2.63e-84
   36    495   15.9    598   6  NGFI_HUMAN  EARLY RESPONSE PROTEI  1.53e-84
   37    492   15.8    462   8  RRXA_HUMAN  RETINOIC ACID RECEPTO  4.52e-84
   38    489   15.7    601   6  NGFI_MOUSE  NUCLEAR HORMONE RECEP  3.95e-83
   39    487   15.6    488   8  RRXA_XENLA  RETINOIC ACID RECEPTO  1.17e-82
   40    473   15.2    598   6  RNR1_RAT    REGENERATING LIVER NU  2.27e-79
   41    469   15.1    467   8  RRXA_CHICK  RETINOIC ACID RECEPTO  1.97e-78
   42    471   15.1    598   6  NOT_HUMAN   IMMEDIATE-EARLY RESPO  6.68e-79
   43    471   15.1    598   7  NUR1_MOUSE  PROBABLE NUCLEAR HORM  6.68e-79
   44    464   14.9    463   8  RRXG_HUMAN  RETINOIC ACID RECEPTO  2.91e-77
   45    464   14.9    463   8  RRXG_MOUSE  RETINOIC ACID RECEPTO  2.91e-77

                              ALIGNMENTS

RESULT     1
ID   ESTR_HUMAN     STANDARD;      PRT;   595 AA.
AC   P03372;
DT   21-JUL-1986 (REL. 01, CREATED)
DT   01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DT   01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE   ESTROGEN RECEPTOR (ER).
GN   ESR.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 86122927.
RA   GREENE G.L., GILNA P., WATERFIELD M., BAKER A., HORT Y., SHINE J.;
RL   SCIENCE 231:1150-1154(1986).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 86146892.
RA   GREEN S., WALTER P., KUMAR V., KRUST A., BORNERT J.-M., ARGOS P.,
RA   CHAMBON P.;
RL   NATURE 320:134-139(1986).
RN   [3]
RP   STRUCTURE BY NMR OF 190-262.
RX   MEDLINE; 91061891.
RA   SCHWABE J.W.R., NEUHAUS D., RHODES D.;
RL   NATURE 348:458-461(1990).
RN   [4]
RP   VARIANT VAL-400.
RX   MEDLINE; 90005402.
RA   TORA L., MULLICK A., METZGER D., PONGLIKITMONGKOL M., PARK I.,
RA   CHAMBON P.;
RL   EMBO J. 8:1981-1986(1989).
RN   [5]
RP   VARIANT ALA-447.
RX   MEDLINE; 92250650.
RA   REESE J.C., KATZENELLENBOGEN B.S.;
RL   J. BIOL. CHEM. 267:9868-9873(1992).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- IN THE ABSENCE OF LIGAND, STEROID HORMONE RECEPTORS ARE THOUGHT
CC       TO BE WEAKLY ASSOCIATED WITH NUCLEAR COMPONENTS; HORMONE BINDING
CC         GREATLY INCREASES RECEPTOR AFFINITY  THE HORMONE-RECEPTOR
CC         COMPLEX APPEARS TO RECOGNIZE DISCRETE DNA SEQUENCES UPSTREAM OF
CC         TRANSCRIPTIONAL START SITES
CC     -!- SUBUNIT: HOMODIMER
CC     -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC         NUCLEAR HORMONE RECEPTORS.
DR   EMBL; X03635; G31234; -.
DR   EMBL; M12674; G182193; -
DR   PIR; A03244; QRHUE.
DR   HSSP; P06536; 1GDC
DR   MIM; 133430; -.
DR   PROSITE; PS00031; STEROID_FINGER
DR   TRANSFAC; T00261; -
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING; POLYMORPHISM.
FT   DOMAIN          1    184       MODULATING.
FT   DNA_BIND      185    250       C4-TYPE ZINC FINGERS (TWO)
FT   ZN_FING       185    205       C4-TYPE.
FT   ZN_FING       221    245       C4-TYPE.
FT   DOMAIN        251    310       HINGE
FT   DOMAIN        311    551       STEROID-BINDING
FT   VARIANT       400    400       G -> V (DESTABILIZES THE RECEPTOR AND
FT                                  DECREASES ITS AFFINITY FOR ESTRADIOL AT
FT                                  25 DEGREES CELSIUS, BUT NOT AT 4 DEGREES
FT                                  CELSIUS).
FT   VARIANT       447    447       G -> A (LOSS OF HORMONE BINDING CAPACITY
FT                                  AND TEMPERATURE-SENSITIVE LOSS IN DNA-
FT                                  BINDING).
SQ   SEQUENCE   595 AA;  66216 MW;  D6945B89 CRC32;

  Query Match             47.3%;  Score 1474;  DB 3;  Length 595;
  Best Local Similarity   57.6%;  Pred. No. 0.00e+00;
  Matches  196;  Conservative   84;  Mismatches   45;  Indels   15;  Gaps   11;

Db     178 saketrycavcndyasgyhygvwscegckaffkrsiqghndymcpatnqctidknrrksc 237
           | :::::||||:|||||||||||||||||||||||||||||||:|||||||||||||||||
Qy      89 SKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSC 148

Db     238 qacrlrkcyevgmmkggirkdrrggrmlkhkrqrddgegrgevgsagdmraanlwpsplm 297
           ||||||||||||||:| | |::| | |:::  |||    | :    :: : :|::     :|
Qy     149 QACRLRKCYEVGMVKCGSRRERCGYRLVR-RQRSADE-QLHCAGKA-KRSGG-HAPR-R 201

Db     298 ikrskknslalsltadqmvsalldaepp-ilyseydptrpfseasmmglltnladrelvh 356
           : |      |   :|::::|:| :||:||||  :| |    |: ||:|||||  ||:|||:||||
Qy     202 V-RELL--LD-ALSPEQLVLTLLEAEPPHVLISR--PSAPFTEASMMMSLTKLADKELVH 255

Db     357 minwakrvpgfvdltlhdqvhllecawleilmiglvwrsmehpgkllfapnllldrnqgk 416
           ||:|||::|:||:|:| |||:|'|  |:|:||:||:|||::||'||:|||:|:||||::||
Qy     256 MISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGK 315

Db     417 cvegmveifdmllatssrfrmmnlqgeefvclksiillnsgvytflsstlksleekdhih 476
           ||||::|||||||||||:||||  ::||  |::|:|::|||||::|  ::::|   :  :     :
Qy     316 CVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSRKLA 374

Db     477 rvldkitdtlihlmakagltlqqqhqrlaqlllilshirh 516
           ::|: :||:|: ::||:|:: |||  ||| ||::|||:||
Qy     375 HLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRH 414

RESULT     2
ID   ESTR_CHICK      STANDARD;      PRT;   589 AA.
AC   P06212;
DT   01-JAN-1988 (REL. 06, CREATED)
DT   01-JAN-1988 (REL. 06, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   ESTROGEN RECEPTOR (ER).
GN   ESR.
OS   GALLUS GALLUS (CHICKEN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;
OC   GALLIFORMES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 86247578.
```

```
RA   KRUST A., GREEN S., ARGOS P., KUMAR V., WALTER P., BORNERT J.-M.,
RA   CHAMBON P.;
RL   EMBO J. 5:891-897(1986).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- IN THE ABSENCE OF LIGAND, STEROID HORMONE RECEPTORS ARE THOUGHT
CC       TO BE WEAKLY ASSOCIATED WITH NUCLEAR COMPONENTS; HORMONE BINDING
CC       GREATLY INCREASES RECEPTOR AFFINITY  THE HORMONE-RECEPTOR
CC       COMPLEX APPEARS TO RECOGNIZE DISCRETE DNA SEQUENCES UPSTREAM OF
CC       TRANSCRIPTIONAL START SITES.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; X03805; G63880; -.
DR   PIR; S07192; S07192.
DR   HSSP; P06536; 1GDC.
DR   PROSITE; PS00031; STEROID_FINGER.
DR   TRANSFAC; T00264; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
FT   DOMAIN          1    178       MODULATING.
FT   DNA_BIND      179    244       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING       179    199       C4-TYPE
FT   ZN_FING       215    239       C4-TYPE.
FT   DOMAIN        245    304       HINGE.
FT   DOMAIN        305    589       STEROID-BINDING.
SQ   SEQUENCE   589 AA;  66746 MW;  85FF1F4A CRC32;

  Query Match             47.2%;  Score 1472;  DB 3;  Length 589;
  Best Local Similarity   51.8%;  Pred. No. 0.00e+00;
  Matches  213;  Conservative   95;  Mismatches   85;  Indels   19;  Gaps   11;

Db     100 spvvflqtapglspfihhhsqqvpyylenepgsfgmreaqppafyrpssdnrrhsirerm 159
           || |:  |: :|||:: |: |     | | : ::       :    ::    : :  :: :
Qy      22 SPNVLWPTPGHLSPLVVHR-QLSHLYAEPQKSPWCEARSLEHTL-PVNRETLKRKVSGNR 79

Db     160 sstnekgslsmestketrycavcndyasgyhygvwscegckaffkrsiqghndymcpatn 219
             ::    |: :    | :::::||||:|||||||||||||||||||'||||||||||||':||||
Qy      80 CASPVTGP-G--SKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATN 136

Db     220 qctidknrrkscqacrlrkcyevgmmkggirkdrrggrmlkhkrqreeqdsrnaeasstc 279
           ||||||'|||||||||||||||||||||:| | |::| |    :::: |   :||     | |  :
Qy     137 QCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSG 196

Db     280 lsaptlwtsplvvkhnkknspalsltadqmvsalleaeppilyseydpsrpfneasmmtl 339
             :|| ::   |::      :  |||   :||:| :||||||| |       |: || |:|||
Qy     197 GHAPRVR-E-LLL-----D--ALS--PEQLVLTLLEAEPPHVLIS-RPSAPFTEASMMMS 244

Db     340 ltnladrelvhminwakrvpgfvdltlhdqvhllecawleilmiglvwrsmehpgkllfa 399
           ||:|||:||||||||:|||::||||:|:| |||:|||   |:|:||:||:|||::||||::||
Qy     245 LTKLADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFA 304

Db     400 pnllldrnqgkcvegmveifdmllataarfrmmnlqgeefvclksiillnsgvytflsst 459
           |:|:|||::|||||||::|||||||||::||| ::||  |::|:|::|||||::| ::::|
Qy     305 PDLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTAT 364

Db     460 lksleerdyihrvldkitdtlihlmaksglslqqqhrrlaqlllilshirh 510
             :   |   | :|: :||:|: ::||:|:: ||| ||| ||::|||:||
Qy     365 QDADSSRKLAH-LLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRH 414

RESULT     3
ID   ESTR_MOUSE      STANDARD;      PRT;   599 AA.
AC   P19785;
DT   01-FEB-1991 (REL. 17, CREATED)
DT   01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
DT   01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE   ESTROGEN RECEPTOR (ER).
GN   ESR OR ESTR.
```

```
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=UTERUS;
RX   MEDLINE; 91042558.
RA   WHITE R., LEES J.A., NEEDHAM M., HAM J., PARKER M.;
RL   MOL. ENDOCRINOL. 1:735-744(1987).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- IN THE ABSENCE OF LIGAND, STEROID HORMONE RECEPTORS ARE THOUGHT
CC       TO BE WEAKLY ASSOCIATED WITH NUCLEAR COMPONENTS; HORMONE BINDING
CC       GREATLY INCREASES RECEPTOR AFFINITY. THE HORMONE-RECEPTOR
CC       COMPLEX APPEARS TO RECOGNIZE DISCRETE DNA SEQUENCES UPSTREAM OF
CC       TRANSCRIPTIONAL START SITES.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; M38651; G193180; -.
DR   PIR; A40061; A40061.
DR   HSSP; P06536; 1GDC.
DR   PROSITE; PS00031; STEROID_FINGER.
DR   TRANSFAC; T00259; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
FT   DOMAIN        1    188       MODULATING.
FT   DNA_BIND    189    254       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     189    209       C4-TYPE.
FT   ZN_FING     225    249       C4-TYPE.
FT   DOMAIN      255    599       STEROID-BINDING.
FT   DOMAIN       64     72       POLY-ALA.
SQ   SEQUENCE   599 AA;  66955 MW;  6BD9EACF CRC32;

  Query Match             46.7%;  Score 1455;  DB 3;  Length 599;
  Best Local Similarity   51.7%;  Pred. No. 0.00e+00;
  Matches  214;  Conservative   93;  Mismatches   87;  Indels   20;  Gaps   14;

Db     109 spsplmllhpppqlspflhphgqqvpyylenepsay-avrdtgppafyrsnsdnrrqngr 166
            : |: :|   | :||: |  | |     | | : |      |              |:  |
Qy      20 TTSPNVLWPTPGHLSP-LVVHRQLSHLYAEPQKSPWCEARSLEHTLPVNRETLKRKVSG- 77

Db     167 erlsssnekgnmimesaketrycavcndyasgyhygvwscegckaffkrsiqghndymcp 226
            :|  :|   |      | :::::|||||:||||||||||||||||||||||||||||:||
Qy      78 NR-CASPVTGP---GSKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICP 133

Db     227 atnqctidknrrkscqacrlrkcyevgmmkggirkdrrggrmlkhkrqrddlegrnemga 286
            ||||||||||||||||||||||||||||:| | |::| | |:::  |||   : :    |
Qy     134 ATNQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRIVR--RQRSA-DEQLH-CA 189

Db     287 sgdmraanlwpsplvikhtkknspalsltadqmvsalldaeppmiyseydpsrpfseasm 346
            :  :|::     :| | :     .  |||  ..|.| .||.||||  .      || ||:||||
Qy     190 GKAKRSGG--HAPRV-RELLLD-ALS--PEQLVLTLLEAEPPHVLIS-RPSAPFTEASM 241

Db     347 mglltnladrelvhminwakrvpgfgdlnlhdqvhllecawleilmiglvwrsmehpgkl 406
            |  ||:|||.|.||||.|||.:|||  .|.| |||  |||   | |.||.|| |||  ||||||
Qy     242 MMSLTKLADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKL 301

Db     407 lfapnllldrnqgkcvegmveifdmllatssrfrmmnlqgeefvclksiillnsgvytfl 466
            :|||:|:||| ::|||||::|||||||||: :||       |  ||  :  :  |||||
Qy     302 IFAPDLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLV 361

Db     467 sstlksleekdhihrvldkitdtlihlmakagltlqqqhrrlaqllllilshirh 520
            ::|  : :     : :.|. .||        :  .  |:  | : |:  ||: .|
Qy     362 TATQDA-DSSRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRH 414


RESULT     4
ID   ESTR_RAT        STANDARD;       PRT;   600 AA.
AC   P06211;
DT   01-JAN-1988 (REL. 06, CREATED)
DT   01-JAN-1988 (REL. 06, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   ESTROGEN RECEPTOR (ER).
GN   ESR OR ESTR.
OS   RATTUS NORVEGICUS (RAT).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WISTAR;
RA   MURAMATSU M.;
RL   SUBMITTED (MAR-1987) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 87174780.
RA   KOIKE S., SAKAI M.;
RL   NUCLEIC ACIDS RES. 15:2499-2513(1987).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=UTERUS;
RA   MAGGI A.M.A.;
RL   SUBMITTED (JUN-1991) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- IN THE ABSENCE OF LIGAND, STEROID HORMONE RECEPTORS ARE THOUGHT
CC       TO BE WEAKLY ASSOCIATED WITH NUCLEAR COMPONENTS; HORMONE BINDING
CC       GREATLY INCREASES RECEPTOR AFFINITY. THE HORMONE-RECEPTOR
CC       COMPLEX APPEARS TO RECOGNIZE DISCRETE DNA SEQUENCES UPSTREAM OF
CC       TRANSCRIPTIONAL START SITES.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; Y00102; G56111; -.
DR   EMBL; X61098; G56121; -.
DR   PIR; S07379; S07379.
DR   HSSP; P06536; 1GDC.
DR   PROSITE; PS00031; STEROID_FINGER.
DR   TRANSFAC; T00258; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
FT   DOMAIN        1    189       MODULATING.
FT   DNA_BIND    190    255       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     190    210       C4-TYPE.
FT   ZN_FING     226    250       C4-TYPE.
FT   DOMAIN      256    600       STEROID-BINDING.
FT   DOMAIN       64     71       POLY-ALA.
FT   CONFLICT    488    488       N -> I (IN REF. 3).
SQ   SEQUENCE   600 AA;  67030 MW;  4CCC59B2 CRC32;

  Query Match             46.5%;  Score 1448;  DB 3;  Length 600;
  Best Local Similarity   51.0%;  Pred. No. 0.00e+00;
  Matches  211;  Conservative   97;  Mismatches   86;  Indels   20;  Gaps   14;

Db     109 spsplmllhppphvspflhphghqvpyylenepsay-avrdtgppafyrsnsdnrrqngr 167
            . || :|   | |:|| |  | :     | | : |:      |              :: :|: :|
Qy      20 TISPNVLWPTPGHLSP-LVVHRQLSHLYAEPQKSPWCEARSLEHTLPVNRETLKRKVSG- 77

Db     168 erlssssekgnmimesaketrycavcndyasgyhygvwscegckaffkrsiqghndymcp 227
            .|  :|.  |     | ...::||||:||||||||||||||||||||||||||||:||
Qy      78 NR-CASPVTGP---GSKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICP 133

Db     228 atnqctidknrrkscqacrlrkcyevgmmkggirkdrrggrmlkhkrqrddlegrnemgt 287
            ||||||||||||||||||||||||||||:| | |::| | |:::: |  |: : :    |
Qy     134 ATNQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRIVRRQRSADE-QLHCA-GK 191

Db     288 sgdmraanlwpsplvikhtkknspalsltadqmvsalldaeppliyseydpsrpfseasm 347
            :  :|::     :| | :     :  |||  ::|:| :||:||||  :      || ||:.||
Qy     192 A--KRSGG--HAPRV-RELLLD--ALS--PEQLVLTLLEAEPPHVLIS-RPSAPFTEASM 241

Db     348 mqlltnladrelvhminwakrvpgfgdlnlhdqvhllecawleilmiglvwrsmehpgkl 407
           |   ||:|:|:|||||:|||::|||  :|:|  |||:|||   |:|.||.||.|||..||||
Qy     242 MMSLTKLADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKL 301

Db     408 lfapnllldrnqgkcvegmveifdmllatssrfrmmnlqgeefvclksiillnsgvytfl 467
           :|||:|:|||::||||||::|||||||||:||||  ::||   |::|:|::|||||::|  ::
Qy     302 IFAPDLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLV 361

Db     468 sstlksleekdhihrvldkindtlihlmakaqltlqqqhrrlaqlllilshirh 521
           ::|  :  :      :  ::|:  :  |:|:  ::||:|::  |||   |||  ||::|||:||
Qy     362 TATQDA-DSSRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRH 414

RESULT     5
ID   ESTR_ORYLA     STANDARD;      PRT;   575 AA.
AC   P50241;
DT   01-OCT-1996 (REL. 34, CREATED)
DT   01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   ESTROGEN RECEPTOR (ER).
GN   ESR.
OS   ORYZIAS LATIPES (MEDAKA FISH).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; PISCES; GNATHOSTOMATA;
OC   OSTEICHTHYES; ACTINOPTERYGII; ATHERINIFORMES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RA   OKADA H.;
RL   SUBMITTED (APR-1994) TO EMBL/GENBANK/DDBJ DATA BANKS
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; D28954; G602326; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
FT   DOMAIN        1    140       MODULATING.
FT   DNA_BIND    141    206       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     141    161       C4-TYPE.
FT   ZN_FING     177    201       C4-TYPE.
FT   DOMAIN      207    575       STEROID-BINDING.
FT   DOMAIN      254    262       POLY-GLY.
FT   DOMAIN      275    278       POLY-LEU.
FT   DOMAIN      466    470       POLY-LEU.
FT   DOMAIN      531    534       POLY-GLY.
SQ   SEQUENCE   575 AA;  62497 MW;  13EE43DC CRC32;

  Query Match             44.3%;  Score 1381;  DB 3;  Length 575;
  Best Local Similarity   52.1%;  Pred. No. 8.70e-303;
  Matches  223;  Conservative   83;  Mismatches   99;  Indels   23;  Gaps   16;

Db      52 ppseqslqs lqsqptsplvfvpsspilspfm-hpp-shhylettstpvyrsshqgasre 108
           |::   :|::    |      |||  |: |::  :||||::  |     || |  |      :|       :        :
Qy       6 PSNVTNLFGGPGRQTTSPNVLWPIPGHLSPLVVHRQLSHLYAEPQKSPWCEARSLEHILP 65

Db     109 dqcisredtesiqelqaqaqaqqfemakdtrfcavcsdyasgyhygvwscegckaffkrs 168
                 :        | |:   |:: :|        :  :|::|||||||||||||||||||||||||||
Qy      66 VNRETLKRKVS-GNRCASPVIGPGSK-RDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRS 123

Db     169 iqghndymcpatnqctidrnrrksqqacrlrkcyevgmmkggvrkdririlrrdkrrtgv 228
           |||||||:||||||||||:||||||||||||||||||||:|  |  |::|       |   ||
Qy     124 IQGHNDYICPATNQCTIDKNRRKSCQACRLRKCYEVGMVKGGSRRERCGY-RLV-RR--- 178

Db     229 pdqlrvvkqqqelkkwvlqqkrsstqqqqqqggrlsvtslppeqvlllqqaepp-ilc 287
              ::     ::|:   :|    |:  :  |  ||  :        |  :  :::||||::|  |    ||||  :|
Qy     179 --QF--SAPEQ--LHCAGKAKRS-GGHAPRVRELL-LDALSPEQLVLTLLEAEPPHVLI 229
```

```
Db     288 srqklsrpytevtmmtlltsmadkelvhmiawakklpgflqlslhdqvllleswlevlm 347
           ||    : |:||  :||   || :||||||||||:|||||:||||:||||  |||  ||||  |:||||
Qy     230 SRPS-A-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLM 287

Db     348 iglıwrsihcpgklifaqdllldrnegdcvegmteifdmllatasrfrvlklkpeefvcl 407
           :||:||||   |||||||  ||:||||:||  ||||:  |||||||||||:||||  |||      |::|:
Qy     288 MGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCV 347

Db     408 kailllnsgafsfctqtmeplhnsaavqsmldtitdallhylsqsqylaqeqarrquqll 467
           ||:||||||: :::  ||:|  ::   :|    :    :|:::||||:    |:  ||   :|:|:  |  |  ||
Qy     348 KAMILLNSSMYPLVTATQDA-DSSRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLL 406

Db     468 lllshirh 475
           :|||||:||
Qy     407 MLLSHVRH 414

RESULT     6
ID   ESTR_SALIR     STANDARD;      PRT;   574 AA.
AC   P16058;
DT   01-APR-1990 (REL. 14, CREATED)
DT   01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   ESTROGEN RECEPTOR (ER).
GN   ESR.
OS   SALMO IRIDEUS (RAINBOW TROUT).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; PISCES; GNATHOSTOMATA;
OC   OSTEICHTHYES; ACTINOPTERYGII; SALMONIFORMES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 91006824.
RA   PAKDEL F., LE GAC F., LE GOFF P., VALOTAIRE Y.;
RL   MOL. CELL. ENDOCRINOL. 71:195-204(1990).
RN   [2]
RP   SEQUENCE OF 150-574 FROM N.A.
RC   TISSUE=LIVER;
RX   MEDLINE; 89127284.
RA   PAKDEL F., LE GUELLEC C., VAILLANT C., LE ROUX M.G., VALOTAIRE Y.;
RL   MOL. ENDOCRINOL. 3:44-51(1989).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- IN THE ABSENCE OF LIGAND, STEROID HORMONE RECEPTORS ARE THOUGHT
CC       TO BE WEAKLY ASSOCIATED WITH NUCLEAR COMPONENTS; HORMONE BINDING
CC       GREATLY INCREASES RECEPTOR AFFINITY. THE HORMONE-RECEPTOR
CC       COMPLEX APPEARS TO RECOGNIZE DISCRETE DNA SEQUENCES UPSTREAM OF
CC       TRANSCRIPTIONAL START SITES.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; M31559; G213784; -.
DR   PIR; A37197; A37197.
DR   PIR; A40070; A40070.
DR   HSSP; P06536; 1GDC.
DR   PROSITE; PS00031; STEROID_FINGER.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
FT   DOMAIN        1    146       MODULATING (POTENTIAL).
FT   DNA_BIND    147    212       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     147    167       C4-TYPE.
FT   ZN_FING     183    207       C4-TYPE.
FT   DOMAIN      213    574       STEROID-BINDING (POTENTIAL).
SQ   SEQUENCE   574 AA;  62872 MW;  B391ADC2 CRC32;

  Query Match             42.6%;  Score 1327;  DB 3;  Length 574;
  Best Local Similarity   57.5%;  Pred. No. 3.32e-289;
  Matches  196;  Conservative   64;  Mismatches   61;  Indels   20;  Gaps   15;

Db     143 etryc avcsdfasgyhygvwscegckaffkrsiqghndymcpatnqctmdrnrrkscqac 202
```

```
              ::::||||||:||||||||||||||||||||||||||||||:|||||||||:|:|||||||||
Qy       92 DAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQAC 151

Db      202 rlrkcyevgmvkggirkdrrg-rvlrkdkryggpagdrekpyhwstqqrpqdgqrnss 261
            ||||||||||| | |::| | |::|:. |  .|| | |   : : : |: ||: ::
Qy      152 FLRKCYEVGMVKCGSRRERCGYRLVRRQ-R---SA-D-EQLHCA-GKAKRS--GGH--AP 200

Db      262 slnggggwtgpritmrpeqvlfllqgatp-alcsrqkvarpyteitmmtllltsmadkelv 320
             :          : :: | : || ||::: |  | ||   | |:|| :||  || :||||||
Qy      201 RVREL--LLDA-LS-PEQLVLTLLEAEPPHVLISRPS-A-PFTEASMMMSLTKLADKELV 254

Db      321 hmiawakkvpgfqelslhdqvqllesswlevlmiglIwrsihcpgklifaqdlildrseg 380
             |:||.|:||' |||| |||:|||| |:||||:||:||||  ||||||| ||:|||' ||
Qy      255 HMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEG 314

Db      381 dcvegmaeifdmllatvsrfgmlklkpeefvclkailllnpgafsfcsnsvqslhnssav 440
             '|||: |||||||| ||| |||   |:.|.||.||||.. ... . . .: :|  .
Qy      315 KCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSRKL 373

Db      441 esmldnitdalihhishsgasvqqqprrqaqllllshirh 481
              :|: :||||:  |: || | |||: | | ||:||||:||
Qy      374 AHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRH 414


RESULT      7
ID   ESTR_SALSA      STANDARD;      PRT,   535 AA.
AC   P50242;
DT   01-OCT-1996 (REL. 34, CREATED)
DT   01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   ESTROGEN RECEPTOR (ER) (FRAGMENT)
GN   ESR.
OS   SALMO SALAR (ATLANTIC SALMON).
OC   EUKARYOTA  METAZOA; CHORDATA; VERTEBRATA; PISCES, GNATHOSTOMATA,
OC   OSTEICHTHYES, ACTINOPTERYGII, SALMONIFORMES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RA   ROGERS S A   LLEWELLYN I   RAMSURN V P   SWEENEY G E., WIGHAM T
RL   SUBMITTED (JUL-1995) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC    !- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC        THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC        PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES
CC   -!- SUBCELLULAR LOCATION  NUCLEAR
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN
CC        A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY)
CC   -!- SIMILARITY  TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC        NUCLEAR HORMONE RECEPTORS.
DR   EMBL; X89959; G929902; .
KW   RECEPTOR  TRANSCRIPTION REGULATION  DNA-BINDING  NUCLEAR PROTEIN
KW   ZINC-FINGER; STEROID-BINDING.
FT   NON_TER       1      1
FT   DOMAIN       <1    104       MODULATING (POTENTIAL)
FT   DNA_BIND    105    170       C4-TYPE ZINC FINGERS (TWO)
FT   ZN_FING     105    125       C4-TYPE.
FT   ZN_FING     141    165       C4-TYPE.
FT   DOMAIN      171    535       STEROID-BINDING (POTENTIAL).
SQ   SEQUENCE   535 AA;  58900 MW;  0E71CD0C CRC32;

  Query Match             42.2%,  Score 1314,  DB 3;  Length 535;
  Best Local Similarity   56.9%;  Pred. No. 6.15e-296;
  Matches   195;  Conservative   61;  Mismatches   66;  Indels   21;  Gaps   14;

Db      101 etrycavcsdfasgyhygfwscegckaffkrsiqghndymcpatnqctmdrnrrkscqac 160
            ::::||||||:|||:||| ||||||||||||||||||||||:|||||||||:|:||||||||
Qy       92 DAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQAC 151

Db      161 rlrkcyevgmvkggirkdrrg-rvlrkdkrycgpagdrekpygdlehrtappqdggrnss 219
            |||||||||||| | |::| | |::|:. |   |            . |  : ||: :
Qy      152 FLRKCYEVGMVKCGSRRERCGYRLVRRQ-R---SADE-QLHCAG---K-AK-RSGGHAPR 201

Db      220 ssslsgggwcgpritmrpeqvlfllqgaeppalcsrqkvarpyteitmmtllltsmadke 279
```

```
              |       : :: | : || ||::  |  | ||   | |:|| :||  || :.|||
Qy      202 VRELL---L-DA-LS-PEQLVLTLLEAEPPHVLISRPS-A-PFTEASMMMSLTKLADKE 252

Db      280 lvhmiawakkvpgfqelslhdqvqllesswlevlmiglIwrslhcpgklifaqdlildrs 339
            |||||:|||||:||| |||| |||:|||| |:||||:||:||||  ||||||| |:|:|:|
Qy      253 LVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRD 312

Db      340 egdcvegmaeifdmllatvsrfrmlklkpeefvclkailllnsgafsfcsnsvqslhnss 399
            || ||||: |||||||| |||| |||   |::|:||:|'|||: ::: : : :: :'
Qy      313 EGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSR 371

Db      400 avesmldnitdalihhishsgasvqqqprrqvqllllshirh 442
              :  :|: :||||:  !: || | |||: !   ||:||||:|!
Qy      372 KLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRH 414


RESULT      8
ID   ESTR_OREAU      STANDARD;      PRT;   314 AA.
AC   P50240;
DT   01-OCT-1996 (REL. 34, CREATED)
DT   01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   ESTROGEN RECEPTOR (ER) (FRAGMENT).
GN   ESR.
OS   OREOCHROMIS AUREUS.
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; PISCES; GNATHOSTOMATA;
OC   OSTEICHTHYES; ACTINOPTERYGII; PERCIFORMES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BLOOD;
RX   MEDLINE, 96272302.
RA   TAN N., LAM T., DING J.;
RL   DNA SEQ. 5:359-370(1995).
CC   -!- FUNCTION  THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC        THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC        PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC        A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -!- SIMILARITY  TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC        NUCLEAR HORMONE RECEPTORS.
DR   EMBL; Z46665, E212276; -.
DR   EMBL, Z46666, E212276, JOINED.
DR   EMBL, Z46667, E212276, JOINED.
DR   EMBL; Z46668, E212276, JOINED.
DR   EMBL; Z46669; E212276; JOINED.
KW   RECEPTOR, TRANSCRIPTION REGULATION, DNA-BINDING, NUCLEAR PROTEIN;
KW   ZINC-FINGER, STEROID-BINDING.
FT   NON_TER       1      1
FT   DOMAIN       <1    314       STEROID-BINDING (POTENTIAL).
SQ   SEQUENCE   314 AA;  34872 MW;  0B18C899 CRC32;

  Query Match             28.2%,  Score 879,  DB 3;  Length 314;
  Best Local Similarity   58.4%;  Pred. No. 1.58e-177;
  Matches   118,  Conservative   40;  Mismatches   40;  Indels    4;  Gaps    4;

Db        1 vllllqgaeppilcsrqkmsrpyteitmtllltsmadkelvhmitwakklpgflqlslhd 60
            || ||::  | :| ||   : |:|| ::|  || :||||||||||:||||:||||::||| '
Qy      216 VLTLLEAEPPHVLISRPS-A-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 273

Db       61 qvlllesswlevlmigllwrslqcpgklifeedlildrnegtcvegmaeifdmllatv-r 119
            || |||| |:||||.||.||||. ||||||  ||.|||.|| ||||. |||||||||  |
Qy      274 QVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSR 333

Db      120 frvlklkpeefvclkailllnsgafsfctgtmeplhdsvavqhmldtitdalifhishfg 179
            || |||    |..:|.||.|||||. ... |.| ..    |   . |.|:::||:|:  |:  |
Qy      334 FRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSRKLAHLLNAVTDALVWVIAKSG 392

Db      180 csaqqqsrrqaqllllshirh 201
             |:|||| | | ||:||||:||
Qy      393 ISSQQQSMRLANLLMLLSHVRH 414
```

```
RESULT    9
ID   ERR2_HUMAN      STANDARD;      PRT;   433 AA.
AC   P11475;
DT   01-OCT-1989 (REL. 12, CREATED)
DT   01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)
DT   01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE   STEROID HORMONE RECEPTOR ERR2.
GN   ERR2
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=HEART;
RX   MEDLINE; 88122546.
RA   GIGUERE V., YANG N., SEGUI P., EVANS R.M.;
RL   NATURE 331:91-94(1988).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; X51417; G36611; -.
DR   PIR; B29345; B29345.
DR   HSSP; P06536; 1GDC.
DR   PROSITE; PS00031; STEROID_FINGER.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER.
FT   DNA_BIND    103    168       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     103    123       C4-TYPE.
FT   ZN_FING     139    163       C4-TYPE.
SQ   SEQUENCE   433 AA;  48287 MW;  673BF88F CRC32;

  Query Match             23.6%;  Score 736;  DB 3;  Length 433;
  Best Local Similarity  43.6%;  Pred. No. 1.89e-142;
  Matches  116;  Conservative   55;  Mismatches  84;  Indels  11;  Gaps   9;

Db    101 rlclvcgdiasgyhygvasceackaffkrtiqgnieyscpatnecoitkrrrkscqacrf 160
           ::| ||:| |||||||| |||:||||||||:|||  :| ||||:| | | ||||||||| 
Qy     94 HFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRL 153

Db    161 mkclkvgmlkegvrldrv--rggrqky-krrldsensyylslqisppakkpltkivs-- 214
            ||  |||:| | | :|    |   ||:     :|       :  |   :|  :  |   :|
Qy    154 RKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPPVRELLLDALSPE 213

Db    215 yl l aepdklyam-ppddpe gdlkalttlcdladrelvfliswakhipgfsnltlgd 271
            | |  : |::     : :     |   :     : :|  |||:||| :||||| |||| :|:| |
Qy    214 QLVLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 273

Db    272 qmsllqsawmeilllqlvyrslpyddklayaedyimdeehsrlv-gllelyrailqlvrr 330
          |: ||:| |||:|::|:: ||:      || :| | ::| : :  | |:|||::  :|     |
Qy    274 QVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSP 333

Db    331 ykklkvekeefvmlkalalansdsmy 356
          :: ||::  |:: :||: | || |||
Qy    334 FRELKLQHKEYLCVKAMILLNS SMY 358


RESULT    10
ID   ERR1_HUMAN      STANDARD;      PRT;   521 AA.
AC   P11474;
DT   01-OCT-1989 (REL. 12, CREATED)
DT   01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)
DT   01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE   STEROID HORMONE RECEPTOR ERR1.
GN   ERR1.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=KIDNEY;
RX   MEDLINE; 88122546.
RA   GIGUERE V., YANG N., SEGUI P., EVANS R.M.;
RL   NATURE 331:91-94(1988).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; X51416; G36609; -.
DR   PIR; A29345; A29345.
DR   HSSP; P06536; 1GDC.
DR   PROSITE; PS00031; STEROID_FINGER.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER.
FT   DNA_BIND    176    241       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     176    196       C4-TYPE.
FT   ZN_FING     212    236       C4-TYPE.
SQ   SEQUENCE   521 AA;  55742 MW;  E2E723EA CRC32;

  Query Match             23.3%;  Score 725;  DB 3;  Length 521;
  Best Local Similarity  40.1%;  Pred. No. 9.05e-140;
  Matches  130;  Conservative   67;  Mismatches 117;  Indels  10;  Gaps   7;

Db    174 rlclvcgdvasgyhygvasceackaffkrtiqgsieyscpasneceitkrrrkacqacrf 233
           ::| ||:| |||||||| |||:||||||||:|||  :| |||:|:| | | |||:|||||:
Qy     94 HFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRL 153

Db    234 tkclrvgmlkegvrldrvrggrqkykrrpevdplpfpg p fpagplavaggprktaapv 291
            ||  |||:| | | :| | |   ::|:  : |  :| :  ::|   :          | :
Qy    154 RKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSP 212

Db    292 nalvshllvvepeklyampdpagpdghlpavatlcdlfdrelvvtiswakslpgfsslsl 351
           : ||  ||  ||  :  :: |::|     : : :| | |:|:| ||||| ||||  |||
Qy    213 EQLVLTLLEAEPPHVL-ISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSL 271

Db    352 sdqmsvlqsvwmevlvlgvaqrslplqdelafaedlvld eegaraaglgelgaallqlv 410
            ||: :|:| |||||::|: ||:       | || ||||| :||      |: |:    ||
Qy    272 FDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATT 331

Db    411 rrlqalrlereeyvllkalalansdsvhiedeprlwssceklh--eal---leyeaqra 465
           |:: |:|:: ||: :||: | ||    :     :   |  ||| |  :|:    |    :::
Qy    332 SRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKS 391

Db    466 gpgqgaerrragrllltlpllrqt 489
          | ..  .  : . |:. |. .:.:
Qy    392 GISSQQQSMRLANLLMLLSHVRHA 415


RESULT    11
ID   PRGR_MOUSE      STANDARD;      PRT;   923 AA.
AC   Q00175;
DT   01-DEC-1992 (REL. 24, CREATED)
DT   01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   PROGESTERONE RECEPTOR (PR).
GN   PGR OR PR.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 91299759.
RA   SCHOTT D.R., SHYAMALA G., SCHNEIDER W., PARRY G.;
RL   BIOCHEMISTRY 30:7014-7020(1991).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; M68915; G200472; -.
DR   PIR; A39596; A39596.
DR   HSSP; P06536; 1GDC.
DR   PROSITE; PS00031; STEROID_FINGER.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
FT   DOMAIN          1     556       MODULATING, PRO-RICH.
FT   DNA_BIND      557     622       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING       557     577       C4-TYPE.
FT   ZN_FING       593     617       C4-TYPE.
FT   DOMAIN        671     923       STEROID-BINDING.
FT   DOMAIN        184     188       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
SQ   SEQUENCE   923 AA;  99073 MW;  EEDFF41A CRC32;

  Query Match              21.9%;  Score 681;  DB 7;  Length 923;
  Best Local Similarity   30.8%;  Pred. No. 4.47e-129;
  Matches  123;  Conservative  112;  Mismatches 138;  Indels  27;  Gaps  27;

Db    491 aapptaaqpaiygplg-lnglpqlgy-y-qaavlkdslpqvyppylnylrpdseasqspqy 547
          ::!|  ::   : : |: | : || :    :|  |   :|! |     |:   :   ::   ||
Qy     28 PTPCHLSPLVVHRQLSHLYAEPQKSPWCEARSLEHTLP-VNRETLK-RKVSGNRCASPVT 85

Db    548 gfdslpq-kicllcgdeasgchygvltcgsckvffkramegqhnylcagrndcivdkirr 606
             |  :  :| :':| ||! |||| :| :|| ||||:::|::..|.|.. |.| .:| ||
Qy     86 GPGSFRCAHFIAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRR 145

Db    607 kncpacrlrkccqagmvl-ggrkfk-kfnkvrvmrtldgvalpqsvglpnesqalsqrit 664
          |:| ||||||| : ||| |:|:    ||  |  |  |   |  |  |    |
Qy    146 KSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQRSADE-QL-HCAGKAKRSGGHAPRVR 203

Db    665 fspnqeiqlvp-plinllmsiepdviyaghdntkpdtsssllts lnqlgerqllsvvkws 723
              : | |  |:  |:  ||  :  :    | | :|::  || |:::::|: :: |:
Qy    204 -ELLLDA-LSPEQLVLTLLEAEPPHVLISRPSA-PFTEASMMMSLTKLADKELVHMISWA 260

Db    724 kslpgfrnlhiddqitliqyswmslmvfglgwrsykhvsgqmlyfapdlilnegr-mk 780
          | :||| :| : .|: |:: || ::: || |||  ! :| :: |||||:|: :: : :
Qy    261 KKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVE 319

Db    781 e-lsfyslcltmwqipqefvklqvtheeflcmkvlllln t-l-pleglrsqsqfeemrss 837
             | :: : | :        |  | : | !:||.| ..!||. : ||      .:.    . :
Qy    319 GIIEIFDM-I-L-ATTSRFRELKIQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAH 375

Db    838 yirelikaig-lrqkgvvptsqrfyqltklldslhdlvkq 876
            :  :  |:  :  |: :.. |.  .|..||  |    |..
Qy    376 LLNAVTDALVWVIAKSGISSQQQSMRLANLL-MLLSHVRH 414


RESULT    12
ID   PRGR_HUMAN      STANDARD;      PRT;   933 AA.
AC   P06401;
DT   01-JAN-1988 (REL. 06, CREATED)
DT   01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   PROGESTERONE RECEPTOR (PR).
GN   PGR.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 90229361.
RA   KASTNER P., KRUST A., TURCOTTE B., STROPP U., TORA L., GRONEMEYER H.,
RA   CHAMBON P.
RL   EMBO J. 9:1603-1614(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 87184565.
RA   MISRAHI M., ATGER M., D'AURIOL L., LOOSFELT H., MERIEL C.,
RA   FRIDLANSKY F., GUIOCHON-MANTEL A., GALIBERT F., MILGROM E.;
RL   BIOCHEM. BIOPHYS. RES. COMMUN. 143:740-748(1987).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- ALTERNATIVE PRODUCTS: THERE ARE TWO FORMS OF THE PROGESTERONE
CC       RECEPTOR: FORM A (THAT STARTS WITH MET-1) AND FORM B (THAT STARTS
CC       WITH MET-165).
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; X51730; G35652; -.
DR   EMBL; M15716; G189935; -.
DR   PIR; A03245; QRHUP.
DR   PIR; S09971; S09971.
DR   HSSP; P06536; 1GDC.
DR   MIM; 264080; -.
DR   PROSITE; PS00031; STEROID_FINGER.
DR   TRANSFAC; T00696; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING; PHOSPHORYLATION; ALTERNATIVE SPLICING.
FT   DOMAIN          1     566       MODULATING, PRO-RICH.
FT   DNA_BIND      567     632       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING       567     587       C4-TYPE.
FT   ZN_FING       603     627       C4-TYPE.
FT   DOMAIN        681     933       STEROID-BINDING.
FT   DOMAIN        183     187       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT   MOD_RES        41      41       PHOSPHORYLATION (BY CK).
FT   MOD_RES       227     227       PHOSPHORYLATION (BY CK).
FT   MOD_RES       232     232       PHOSPHORYLATION (BY CK).
FT   MOD_RES       552     552       PHOSPHORYLATION (BY CK).
FT   MOD_RES       797     797       PHOSPHORYLATION (BY CK).
FT   VARSPLIC        1     164       MISSING (IN FORM B).
FT   CONFLICT      220     220       S -> C (IN REF. 2).
FT   CONFLICT      256     256       C -> V (IN REF. 2).
FT   CONFLICT      344     344       I -> S (IN REF. 2).
FT   CONFLICT      660     660       L -> V (IN REF. 2).
SQ   SEQUENCE   933 AA;  98941 MW;  CE1124EB CRC32;

  Query Match              21.4%;  Score 666;  DB 7;  Length 933;
  Best Local Similarity   33.4%;  Pred. No. 1.92e-125;
  Matches  110;  Conservative  89;  Mismatches 113;  Indels  17;  Gaps  17;

Db    566 iclicgdeasgchygvltcgsckvffkramegqhnylcagrndcivdkirrkncpacrlr 625
           :| :|:| ||| |||| :| :|| |||:::|:::|:|:: !:| :|! |||:! !||||
Qy     95 FIAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRLR 154

Db    626 kccqagmvl-ggrkfk-kfnkvrvvraldavalpqpvgvpnesqalsqrftfspgqdiq 683
          || : |||  |:!: :  :  ||  |: |    | :  | :: | : : |        !  |
Qy    155 KCYEVGMVKCGSRRERCGYRLVRRQRSADE-QL-HCAGKAKRSGGHAPRVR-ELLLDA-L 210

Db    684 ip-plinllmsiepdviyaghdntkpdtsssllts lnqlgerqllsvvkwskslpgfrnl 742
           |  |:  |:  ||  :  :: :. | | .|:: ||  |..:.:|. :: |:| :||| :|
Qy    211 SPEQLVLTLLEAEPPHVLISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVEL 269

Db    743 hiddqitliqyswmslmvfglgwrsykhvsgqmlyfapdlilneqrmk-essfyslcltm 801
           : ||: |::  || ::: || |||  | :| :: |||||:|: :  |   ::  :  :
Qy    270 SLFDQVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGIIEIFDML 327

Db    802 wqipqefvklqvsqeeflcmkvllllntipleglrsqtq-feemrs-syi-relikaig- 857
                |  | : : |:||:| ::|||: ::  | : ||   :  |  : :   : |:
Qy    328 LATTSRFRELKLQHKEYLCVKAMILLNS-SMYPLVTATQDADSSRKLAHLLNAVIDALVW 386

Db    858 lrqkgvvsssqrfyqltklldnlhdlvkq 886
           .  |: .:| |:  .|..||  |   |::
Qy    387 VIAKSGISSQQQSMRLANLL-MLLSHVRH 414


RESULT    13
ID   PRGR_RABIT      STANDARD;      PRT;   930 AA.
AC   P06186;
DT   01-JAN-1988 (REL. 06, CREATED)
DT   01-JAN-1988 (REL. 06, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   PROGESTERONE RECEPTOR (PR).
GN   PGR.
OS   ORYCTOLAGUS CUNICULUS (RABBIT).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; LAGOMORPHA.
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RX   MEDLINE; 87067449.
RA   LOOSFELT H., ATGER M., MISRAHI M., GUIOCHON-MANTEL A., MERIEL C.,
RA   LOGEAT F., BENAROUS R., MILGROM E.;
RL   PROC. NATL. ACAD. SCI. U.S.A. 83:9045-9049(1986)
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS
DR   EMBL; M14547, G165632,  .
DR   PIR; A25923; A25923.
DR   HSSP; P06536; 1GDC.
DR   PROSITE; PS00031; STEROID_FINGER.
DR   TRANSFAC; T00697; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
FT   DOMAIN          1    565       MODULATING, PRO RICH.
FT   DNA_BIND      568    633       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING       568    588       C4-TYPE.
FT   ZN_FING       604    628       C4-TYPE.
FT   DOMAIN        678    930       STEROID-BINDING.
FT   DOMAIN        184    188       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL)
SQ   SEQUENCE   930 AA;  98666 MW;  EE133E86 CRC32;

  Query Match              21.2%;  Score 662;  DB 7;  Length 930;
  Best Local Similarity   32.5%;  Pred. No. 1.79e-124;
  Matches  106;  Conservative   91;  Mismatches 114;  Indels  15;  Gaps  15;

Db    567 iclicgdeasgchygvltcgsckvffkramegqhnylcagrndcivdkirrkncpacrlr 626
          :| :|:| ||| |||| :| :|| ||||:::|:::|:|:: |:| :|| |||:| |||||
Qy     95 FCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRLR 154

Db    627 k-ccqagmvl-grkfk-kfnkvrvmraldavalpqpvgipnesq-rltfspsqplq-llp 682
          || : |||  |:|: :  :  ||  |: |   | :  | :: |  :         :: | |
Qy    155 KCYEVGMVKCGSRRERCGYRIVRRQRSADE-QL-HCAGKAKRSGGHAPRVRELLLDAISP 212

Db    683 -plinllmsiepdviyaghdntkpdtsssllts lnqlgerqllsvvkwskslpgfrnlhi 741
            |:  |:  ||  :  :: :: | | :|:: ||  |::::|: :: |:| :||| :| .
Qy    213 EQLVLTLLEAEPPHVLISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSL 271

Db    742 ddqitliqyswmslmvfglgwrsykhvsgqmlyfapdlilneqrmk-essfyslcltmwq 800
           ||: |::  || ::: || |||  | :| :: |||||:|: :  |    :: :    :
Qy    272 HDQVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDMLLA 329

Db    801 ipqefvklqvsqeeflcmkvllllnt-i-pleglrsqsqfeemrssyirelikaig-lrq 857
              |  | : : |:||:| ::||: : ||      :::   : : :  :  |: :
Qy    330 TTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIA 389

Db    858 kgvvsssqrfyqltklldnlhdlvkq 883
          |: :|| |:  :|::||  |   |::
Qy    390 KSGISSQQQSMRLANLL-MLLSHVRH 414


RESULT    14
ID   GCR_HUMAN      STANDARD;      PRT;   777 AA.
AC   P04150; P04151;
DT   01-NOV-1986 (REL. 03, CREATED)
DT   01-NOV-1986 (REL. 03, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   GLUCOCORTICOID RECEPTOR (GR).
GN   GRL.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA, VERTEBRATA, TETRAPODA, MAMMALIA,
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 86092206.
RA   HOLLENBERG S.M., WEINBERGER C., ONG E.S., CERELLI G., ORO A.,
RA   LEBO R., THOMPSON E.B., ROSENFELD M.G., EVANS R.M.;
RL   NATURE 318:635-641(1985).
```

```
RN   [2]
RP   SEQUENCE FROM N.A.
RA   MUNROE D.G., PANG J., TAYLOR G.R., LAU C., PLANTE R.K., ZHOU L.;
RL   SUBMITTED (SEP-1993) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [3]
RP   SEQUENCE OF 1-394 FROM N.A.
RX   MEDLINE; 91224961.
RA   LECLERC S., XIE B.X., ROY R., GOVINDAN M.V.;
RL   J. BIOL. CHEM. 266:8711-8719(1991).
RN   [4]
RP   DOMAINS.
RX   MEDLINE; 86092211.
RA   WEINBERGER C., HOLLENBERG S.M., ROSENFELD M.G., EVANS R.M.;
RL   NATURE 318:670-672(1985).
RN   [5]
RP   VARIANT SER-363.
RX   MEDLINE; 93187003.
RA   KARL M., LAMBERTS S.W.J., DETERA-WADLEIGH S.D., ENCIO I.J.,
RA   STRATAKIS C.A., HURLEY D.M., ACCILI D., CHROUSOS G.P.;
RL   J. CLIN. ENDOCRINOL. METAB. 76:683-689(1993).
RN   [6]
RP   VARIANT VAL-641.
RX   MEDLINE; 91123468.
RA   HURLEY D.M., ACCILI D., STRATAKIS C.A., KARL M., VAMVAKOPOULOS N.,
RA   RORER E., CONSTANTINE K., TAYLOR S.I., CHROUSOS G.P.;
RL   J. CLIN. INVEST. 87:680-686(1991).
RN   [7]
RP   VARIANT ILE-729.
RX   MEDLINE; 93253031.
RA   MALCHOFF D.M., BRUFSKY A., REARDON G., MCDERMOTT P., JAVIER E.C.,
RA   BERGH C.H., ROWE D., MALCHOFF C.D.;
RL   J. CLIN. INVEST. 91:1918-1925(1993).
RN   [8]
RP   VARIANT PHE-753.
RX   MEDLINE; 93302771.
RA   ASHRAF J., THOMPSON E.B.;
RL   MOL. ENDOCRINOL. 7:631-642(1993).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES. BINDS TO THE
CC       GRE TARGET SITE.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
CC   -!- ALTERNATIVE PRODUCTS: ALPHA AND BETA FORMS OF HGR DIFFER ONLY IN
CC       THEIR CARBOXYL-TERMINAL SEQUENCES. THEY ARE GENERATED BY
CC       ALTERNATIVE SPLICING. ALPHA-HGR IS THE PREDOMINANT PHYSIOLOGICAL
CC       FORM OF HGR. THE IN VITRO TRANSLATION PRODUCT OF HGR BETA DOES NOT
CC       BIND STEROID ITS FUNCTION IS UNCLEAR; HOWEVER, IT IS POSSIBLE THAT
CC       VARIANT RECEPTORS PERFORM TISSUE-SPECIFIC FUNCTIONS.
CC   -!- DISEASE: DEFECTS IN GRL ARE A CAUSE OF FAMILIAL GLUCOCORTICOID
CC       RESISTANCE (CORTISOL RESISTANCE); A HYPERTENSIVE, HYPERANDROGENIC
CC       DISORDER CHARACTERIZED BY INCREASED SERUM CORTISOL CONCENTRATIONS.
DR   EMBL; M10901; G183033; -.
DR   EMBL; M11050; G183035; -.
DR   EMBL; U01351; G458657; -.
DR   EMBL; X03225; G31680; -.
DR   EMBL; X03348; G31682; -.
DR   EMBL; M69104; G553322; -.
DR   EMBL; M73816; G553323; -.
DR   EMBL; S68378; G239758; -.
DR   PIR; A03246; QRHUGA.
DR   PIR; A03247; QRHUGB.
DR   HSSP; P06536; 1GDC.
DR   MIM; 138040; -.
DR   PROSITE; PS00031; STEROID_FINGER.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING; DISEASE MUTATION; POLYMORPHISM;
KW   ALTERNATIVE SPLICING.
FT   DOMAIN        145    278       IMMUNOGENIC.
FT   DNA_BIND      421    486       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     421    441       C4-TYPE.
FT   ZN_FING     457    481       C4-TYPE.
FT   DOMAIN      487    527       HINGE.
FT   DOMAIN      528    777       STEROID-BINDING.
FT   VARIANT     363    363       N -> S.
FT   VARIANT     641    641       D -> V (IN GLUCOCORTICOID RESISTANCE).
FT   VARIANT     729    729       V -> I (IN GLUCOCORTICOID RESISTANCE).
FT   VARIANT     753    753       L -> F (IN GLUCOCORTICOID RESISTANCE).
FT   VARSPLIC    728    777       VVENLLNYCFQTFLDKTMSIEFPEMLAEIITNQIPKYSNGN
FT                                IKKLLFHQK -> ENVMWLKPESTSHTLI (IN BETA
FT                                FORM).
FT   CONFLICT    451    451       G -> GR (IN REF. 2).
SQ   SEQUENCE   777 AA;  85659 MW;  D33D1F01 CRC32;

  Query Match             19.7%;  Score 614;  DB 4;  Length 777;
  Best Local Similarity   33.4%;  Pred. No. 6.84e-113;
  Matches   97;  Conservative   84;  Mismatches   93;  Indels   16;  Gaps   15;

Db    397 smrpdvssppsssstattgp-p-pklclvcsdeasgchygvltcgsckvffkravegqhn 454
            :::  ||:   :|: :  |:     : :| |||| ||| ||||  |  :|| ||||:  :
Qy     70 TIKEKVSGNRCASPVIGPGSKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHND 129

Db    455 ylcagrndciidkirrkncpacryrkclqagmnl-earktk-k-kikgiqqattg-v-sq 509
          |:|:: |.| ||| |||.| ||| ||| : ||     .|. .    ..   |...  .
Qy    130 YICPATNQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCA 189

Db    510 etsenpgnkt-ivpatlpq-ltpt-lv-slleviepevlyagydssvpdstwrimttlnm 565
             :  :|  :  :    | : |.|  ||  |||   : ::          .    :  :
Qy    190 GKAKRSGGHAPRVRELLLDALSPEQLVLTLLEAEPPHVLISRPSAPFTEAS--MMMSLTK 247

Db    566 lggrqviaavkwakaipgfrnlhlddqmtllqyswmflmafalgwrsyrqssanllcfap 625
          |: ::::   : ||| |||| .| | ||. ||.  ||      | |||       :|: |||
Qy    248 LADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPG-KLI-FAP 305

Db    626 dliin-eqrmtlpcmydqckhmlyvsselhrlqvsyeeylcmktllllss 674
          ||:..  .  .  .   .   .   :|  .|  .  .  .   |||| |   ||
Qy    306 DLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNS 355


RESULT   15
ID   GCR_XENLA      STANDARD;      PRT;   776 AA.
AC   P49844;
DT   01-OCT-1996 (REL. 34, CREATED)
DT   01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   GLUCOCORTICOID RECEPTOR (GR).
OS   XENOPUS LAEVIS (AFRICAN CLAWED FROG).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AMPHIBIA; ANURA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=EMBRYO;
RX   MEDLINE; 94289478.
RA   GAO X., KALKHOVEN E., PETERSON-MADURO J., VAN DER BURG B.,
RA   DESTREE O.H.J.;
RL   BIOCHIM. BIOPHYS. ACTA 1218:194-198(1994).
RN   [2]
RP   SEQUENCE OF 452-776 FROM N.A.
RC   TISSUE=LIVER;
RA   PICARD D.;
RL   SUBMITTED (FEB-1994) TO EMBL/GENBANK/DDBJ DATA BANKS
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES. BINDS TO THE
CC       GRE TARGET SITE.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOID
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; X72211; G444043; -.
DR   EMBL; X77764; G456374; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
```

```
FT   DOMAIN        1    419       MODULATING.
FT   DNA_BIND    420    485       C4 TYPE ZINC FINGERS (TWO).
FT   ZN_FING     420    440       C4-TYPE.
FT   ZN_FING     456    480       C4-TYPE.
FT   DOMAIN      486    776       STEROID-BINDING.
FT   CONFLICT    502    502       S -> P (IN REF. 2).
FT   CONFLICT    504    504       T -> A (IN REF. 2).
FT   CONFLICT    506    506       T -> A (IN REF. 2).
FT   CONFLICT    511    511       P -> N (IN REF. 2).
FT   CONFLICT    519    519       L -> M (IN REF. 2).
FT   CONFLICT    544    544       Y -> F (IN REF. 2).
FT   CONFLICT    551    551       I -> M (IN REF. 2).
FT   CONFLICT    654    654       S -> R (IN REF. 2).
FT   CONFLICT    740    740       L -> M (IN REF. 2).
FT   CONFLICT    768    768       L -> I (IN REF. 2).
SQ   SEQUENCE   776 AA;  84977 MW;  895C496D CRC32;

  Query Match             19.5%;  Score 608;  DB 4;  Length 776;
  Best Local Similarity   34.9%;  Pred. No. 1.98e-111;
  Matches   99;  Conservative   74;  Mismatches   96;  Indels   15;  Gaps   15;

Db    398 irsdaspsrstsststgpppklclvcsdeasgchygvltcgsckvffkravegqhnylca 457
           . .:   .|| |..  '      . .| |||| ||| |||| .| :|| ||||::: ::::|:!:
Qy     75 VSGNRCASPVIGPGSK-RDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICP 133

Db    458 grndciidkirrkncpacryrkclqagmnl-earktk-k-kikgiqqsttatar-es-pe 512
          : |.| ||| |||:| ||| ||| : ||     :|: :     ::    |:|:      :  :  :
Qy    134 ATNQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAK 193

Db    513 tsmtrtl-vpasv-aqltpt-li-sllevieperlysgydssipdttrrlmsslnmlggr 568
           |  ..  |    .    |.|  |. .|||    | || |     .:: :::   :|| | :': :
Qy    194 RSGGHAPRVRELLLDALSPEQLVLTLLEAEPPHVLISRPSAPFTEASM-MMS-LTKLADK 251

Db    569 qvvsavrwakaipgfrnlhlddqmtllqyswmflmvfalgwrsykqtnasllyfapdlvi 628
           .:|   : ||| |||| :| | ||: ||:   || ::: :| |||   :   |   | || .||:
Qy    252 ELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHP-GK-LIFAPDLVL 309

Db    629 t-edrmhlpfmqercqgmlkiagcmsslqisydeylcmkvlllm 671
            :: :  :   :  |   :  :!  ::     |  :    ||||:| ::|:
Qy    310 DRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILL 353

Search completed. Thu Dec 18 09:07:50 1997
Job time : 64 secs.
```

```
**********************************************************************

                          MPSRCH                              (TM)

**********************************************************************

           Release 2 1D John F. Collins, Biocomputing Research Unit
           Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
                  Distribution rights by IntelliGenetics, Inc.

MPsrch_pp   protein - protein database search, using Smith-Waterman algorithm

Run on:            Thu Dec 19 09:16:50 1997;  MasPar time 13.36 Seconds
                                              841.147 Million cell updates/sec
Tabular output not generated.

Title:             >US-08-826-361A-25
Description:       (1-530) from US08826361A.pep
Perfect Score:     3904
Sequence:          1 MDIKNSPSSLNSPSSYNCSQ. . . . .    ECSPAEDSKSKEGSQNPQSQ 530

Scoring table:     PAM 150
                   Gap 11

Searched:          59021 seqs, 21210388 residues

Post-processing:   Minimum Match 0%
                   Listing first 45 summaries

Database:          swiss-prot34
                     1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
                     8:part8 9:part9 10:part10 11:part11

Statistics:        Mean 50.591;  Variance 84.712;  scale 0.597

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution

                                SUMMARIES
                    %
Result            Query
   No.  Score     Match Length DB  ID           Description            Pred. No.
-------------------------------------------------------------------------------
     1   1696      43.5     589  3  ESTR_CHICK   ESTROGEN RECEPTOR (ER   0.00e+00
     2   1693      43.4     595  3  ESTR_HUMAN   ESTROGEN RECEPTOR (ER   0.00e+00
     3   1669      42.8     599  3  ESTR_MOUSE   ESTROGEN RECEPTOR (ER   0.00e+00
     4   1663      42.6     600  3  ESTR_RAT     ESTROGEN RECEPTOR (ER   0.00e+00
     5   1591      40.8     575  3  ESTR_ORYLA   ESTROGEN RECEPTOR (ER   0.00e+00
     6   1516      38.8     574  3  ESTR_SALIR   ESTROGEN RECEPTOR (ER   0.00e+00
     7   1503      38.5     535  3  ESTR_SALSA   ESTROGEN RECEPTOR (ER   0.00e+00
     8   1076      27.6     314  3  ESTR_TREAU   ESTROGEN RECEPTOR (ER   2.03e-224
     9    807      20.7     433  3  ERR2_HUMAN   STEROID HORMONE RECEP   4.06e-158
    10    800      20.5     521  3  ERR1_HUMAN   STEROID HORMONE RECEP   2.16e-155
    11    681      17.4     923  7  PRGR_MOUSE   PROGESTERONE RECEPTOR   1.39e-127
    12    666      17.1     933  7  PRGR_HUMAN   PROGESTERONE RECEPTOR   5.61e-124
    13    662      17.0     930  7  PRGR_RABIT   PROGESTERONE RECEPTOR   5.13e-123
    14    614      15.7     777  4  GCR_HUMAN    GLUCOCORTICOID RECEPT   1.61e-111
    15    608      15.6     776  4  GCR_XENLA    GLUCOCORTICOID RECEPT   4.36e-110
    16    597      15.3     786  7  PRGR_CHICK   PROGESTERONE RECEPTOR   1.83e-107
    17    586      15.0     783  4  GCR_MOUSE    GLUCOCORTICOID RECEPT   7.57e-105
    18    579      14.8     133  3  ESTR_BOVIN   ESTROGEN RECEPTOR (ER   3.49e-103
    19    577      14.8     771  4  GCR_CAVPO    GLUCOCORTICOID RECEPT   1.04e-102
    20    577      14.8     795  4  GCR_RAT      GLUCOCORTICOID RECEPT   1.04e-102
    21    571      14.6     899  1  ANDR_MOUSE   ANDROGEN RECEPTOR.      2.76e-101
    22    570      14.6     902  1  ANDR_RAT     ANDROGEN RECEPTOR.      4.77e-101
    23    568      14.5     709  1  ANDR_RABIT   ANDROGEN RECEPTOR (FR   1.42e-100
    24    567      14.5     919  1  ANDR_HUMAN   ANDROGEN RECEPTOR.      2.46e-100
    25    562      14.4     981  6  MCR_RAT      MINERALOCORTICOID REC   3.76e-99
    26    544      13.9     758  4  GCR_ONCMY    GLUCOCORTICOID RECEPT   1.79e-95
    27    535      13.7     533  8  RRXB_HUMAN   RETINOIC ACID RECEPTO   1.67e-92
    28    529      13.6     458  8  RRXB_RAT     RETINOIC ACID RECEPTO   2.34e-91
    29    529      13.6     520  8  RRXB_MOUSE   RETINOIC ACID RECEPTO   2.34e-91
    30    523      13.4     984  6  MCR_HUMAN    MINERALOCORTICOID REC   6.04e-90
    31    517      13.2     111  3  ESTR_SHEEP   ESTROGEN RECEPTOR (ER   1.55e-88
    32    514      13.2     467  8  RRXA_MOUSE   RETINOIC ACID RECEPTO   7.87e-88
    33    514      13.2     467  8  RRXA_RAT     RETINOIC ACID RECEPTO   7.87e-88
    34    513      13.1     462  8  RRXA_HUMAN   RETINOIC ACID RECEPTO   1.35e-87
    35    507      13.0     488  8  RRXA_XENLA   RETINOIC ACID RECEPTO   3.45e-86
    36    502      12.9     597  6  NGFI_RAT     NERVE GROWTH FACTOR I   5.12e-85
    37    500      12.9     598  6  NGFI_JANFA   ORPHAN NUCLEAR RECEPT   2.99e-85
    38    499      12.8     121  3  ESTR_MACMU   ESTROGEN RECEPTOR (ER   2.58e-84
    39    501      12.8     598  6  NGFI_HUMAN   EARLY RESPONSE PROTEI   8.79e-85
    40    493      12.6     467  8  RRXA_CHICK   RETINOIC ACID RECEPTO   6.53e-83
    41    491      12.6     601  6  NGFI_MOUSE   NUCLEAR HORMONE RECEP   1.92e-82
    42    489      12.5     463  8  RRXG_MOUSE   RETINOIC ACID RECEPTO   5.62e-82
    43    488      12.5     463  8  RRXG_HUMAN   RETINOIC ACID RECEPTO   9.59e-82
    44    474      12.1     577  6  NGFI_XENLA   NERVE GROWTH FACTOR I   1.75e-78
    45    473      12.1     598  8  RNR1_RAT     REGENERATING LIVER NU   3.01e-78

                                ALIGNMENTS

RESULT     1
ID   ESTR_CHICK     STANDARD;      PRT;   589 AA.
AC   P06212;
DT   01-JAN-1988 (REL. 06, CREATED)
DT   01-JAN-1988 (REL. 06, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   ESTROGEN RECEPTOR (ER).
GN   ESR.
OS   GALLUS GALLUS (CHICKEN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;
OC   GALLIFORMES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 86247578.
RA   KRUST A., GREEN S., ARGOS P., KUMAR V., WALTER P., BORNERT J.-M.,
RA   CHAMBON P.;
RL   EMBO J. 5:891-897(1986).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID BINDING DOMAIN.
CC   -!- IN THE ABSENCE OF LIGAND, STEROID HORMONE RECEPTORS ARE THOUGHT
CC       TO BE WEAKLY ASSOCIATED WITH NUCLEAR COMPONENTS; HORMONE BINDING
CC       GREATLY INCREASES RECEPTOR AFFINITY. THE HORMONE RECEPTOR
CC       COMPLEX APPEARS TO RECOGNIZE DISCRETE DNA SEQUENCES UPSTREAM OF
CC       TRANSCRIPTIONAL START SITES.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; X03805; G63380; -.
DR   PIR; S07192; S07192.
DR   HSSP; P03372; 1HCQ.
DR   PROSITE; PS00031; STEROID_FINGER.
DR   TRANSFAC; T00264; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
FT   DOMAIN          1    178       MODULATING.
FT   DNA_BIND      179    244       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING       179    199       C4-TYPE.
FT   ZN_FING       215    239       C4-TYPE.
FT   DOMAIN        245    304       HINGE.
FT   DOMAIN        305    589       STEROID-BINDING.
SQ   SEQUENCE   589 AA;  66740 MW;  3CFF1F4A CRC32;
```

```
  Query Match             43.5%;  Score 1698;  DB 3;  Length 589;
  Best Local Similarity   51.4%;  Pred. No. 0.00e+00;
  Matches  246;  Conservative  112;  Mismatches  100;  Indels   21;  Gaps   14;

Db      69 yrsttlsyaptsesfq-ssslagths lnsvppspvvflqtapqlspflhhhsqqvpyyl 126
           |  : |: |:|:   ::: :|::: :  :          || |:   |: :|||:: |: |      |
Qy      43 YPAMTF-YSPAVMNYSIPSNVTNLEGGPGRQTTSPNVLWPTPGHLSPLVVHR-QLSHLYA 100

Db     127 enpqqsfqpreaappafyrpssdnrrhsirermsstnekgslsmestketrycavcndya 186
           | : ::       :  ::  : :  :: :       ::    |: :   | :::::||||:|||
Qy     101 EPQKSPWCEARSLEHTL-PVNRETLKRKVSGNRCASPVTGP-G--SKRDAHFCAVCSDYA 156

Db     187 sgyhygvwscegckaffkrsiqghndymcpatnqctidknrrkscqacrlrkcyevgmmk 246
           |||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||:|
Qy     157 SGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRKCYEVGMVK 216

Db     247 ggirkdrrggqmmkqkrqreeqdsrngeasstelraptlwtsplvvkhnkknspalslta 306
            | |::| :  :::: |  :||      | |   :   :|| ::    |::        :  ||| :
Qy     217 CGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVR-E-LLL----D-ALS--P 265

Db     307 eqmvsalleaepplvyseydptrpfneasmmtlltnladrelvhminwakrvpgfvdltl 366
           ||:| :|||||||| |       |: || |||||   ||:|||:||||||||:|||::|||||:|:|
Qy     266 EQLVLTLLEAEPPHVLIS-RPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSL 324

Db     367 hdqvhllecawleilmiglvwrsmehpgkllfapnllldrnqgkcvegmveifdmllats 426
            |||:|||   |:|:||:||:||::|||||:|||:|:|||::||||||::||||||||:
Qy     325 FDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATT 384

Db     427 srfrmmnlqgeefvclksiillnsgvytflsstlksleerdyihrvldkitdtlihlmak 486
           :||| ::||  |::|:|::||||||::| ::::|  :     |     | :|: :||:|: ::||
Qy     385 SRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAH-LLNAVTDALVWVIAK 443

Db     487 sglslqqqhrrlaqllliilshirhmsnkgmehlynmkcknvvplydlllemldahrlha 545
           ||:| |||  ||| ||::|||:|  |||||||| ||||||||||:|||||||||:|| |::
Qy     444 SGISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNAHVLRG 502


RESULT     2
ID   ESTR_HUMAN      STANDARD;      PRT;   595 AA.
AC   P03372;
DT   21-JUL-1986 (REL. 01, CREATED)
DT   01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DT   01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE   ESTROGEN RECEPTOR (ER).
GN   ESR.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 86122927.
RA   GREENE G.L., GILNA P., WATERFIELD M., BAKER A., HORT Y., SHINE J.;
RL   SCIENCE 231:1150-1154(1986).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 86146892.
RA   GREEN S., WALTER P., KUMAR V., KRUST A., BORNERT J.-M., ARGOS P.,
RA   CHAMBON P.;
RL   NATURE 320:134-139(1986).
RN   [3]
RP   STRUCTURE BY NMR OF 180-262.
RX   MEDLINE; 91061891.
RA   SCHWABE J.W.R., NEUHAUS D., RHODES D.;
RL   NATURE 348:458-461(1990).
RN   [4]
RP   VARIANT VAL-400.
RX   MEDLINE; 90005402.
RA   TORA L., MULLICK A., METZGER D., PONGLIKITMONGKOL M., PARK I.,
RA   CHAMBON P.;
RL   EMBO J. 8:1981-1986(1989).
RN   [5]
RP   VARIANT ALA-447.
RX   MEDLINE; 92250650.
RA   REESE J.C., KATZENELLENBOGEN B.S.;
RL   J. BIOL. CHEM. 267:9868-9873(1992).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- IN THE ABSENCE OF LIGAND, STEROID HORMONE RECEPTORS ARE THOUGHT
CC       TO BE WEAKLY ASSOCIATED WITH NUCLEAR COMPONENTS; HORMONE BINDING
CC       GREATLY INCREASES RECEPTOR AFFINITY. THE HORMONE-RECEPTOR
CC       COMPLEX APPEARS TO RECOGNIZE DISCRETE DNA SEQUENCES UPSTREAM OF
CC       TRANSCRIPTIONAL START SITES.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; X03635; G31234; -.
DR   EMBL; M12674; G182193; -.
DR   PIR; A03244; QRHUE.
DR   HSSP; P06536; 1GDC.
DR   MIM; 133430; -.
DR   PROSITE; PS00031; STEROID_FINGER.
DR   TRANSFAC; T00261; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING; POLYMORPHISM.
FT   DOMAIN          1    184       MODULATING.
FT   DNA_BIND      185    250       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING       185    205       C4-TYPE.
FT   ZN_FING       221    245       C4-TYPE.
FT   DOMAIN        251    310       HINGE.
FT   DOMAIN        311    551       STEROID-BINDING.
FT   VARIANT       400    400       G -> V (DESTABILIZES THE RECEPTOR AND
FT                                  DECREASES ITS AFFINITY FOR ESTRADIOL AT
FT                                  25 DEGREES CELSIUS, BUT NOT AT 4 DEGREES
FT                                  CELSIUS).
FT   VARIANT       447    447       C -> A (LOSS OF HORMONE BINDING CAPACITY
FT                                  AND TEMPERATURE-SENSITIVE LOSS IN DNA-
FT                                  BINDING).
SQ   SEQUENCE   595 AA;  66216 MW;  D6945BB9 CRC32;

  Query Match             43.4%;  Score 1693;  DB 3;  Length 595;
  Best Local Similarity   56.8%;  Pred. No. 0.00e+00;
  Matches  229;  Conservative   98;  Mismatches   59;  Indels   17;  Gaps   13;

Db     178 saketrycavcndyasgyhygvwscegckaffkrsiqghndymcpatnqctidknrrksc 237
           | :::::||||:|||||||||||||||||||||||||||||||||:||||||||||||||||||
Qy     142 SKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSC 201

Db     238 qacrlrkcyevgmmkggirkdrrggrmlkhkrqrddgegrgevgsagdmraanlwpsplm 297
           |||||||||||||:| | |::| | |:::  |||   | :   :: : :|::      :|
Qy     202 QACRLRKCYEVGMVKCGSRRERCGYRLVR--RQRSADE-QLHCAGKA-KRSGG--HAP-R 254

Db     298 ikrskknslalsltadqmvsalldaepp ilyseydptrpfseasmmglltnladrelvh 356
           : |      |   :|::::|:| :||:||||  :| |     |: ||:|||||  ||:||||:||||
Qy     255 V-RELL--LD-ALSPEQLVLTLLEAEPPHVLISR--PSAPFTEASMMMSLTKLADKELVH 308

Db     357 minwakrvpgfvdltlhdqvhllecawleilmiglvwrsmehpgkllfapnllldrnqgk 416
           ||:|||::||||:|:| |||:|||  |:|:||:||:||::|||||:|||:|:|||::||
Qy     309 MISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGK 368

Db     417 cvegmveifdmllatssrfrmmnlqgeefvclksiillnsgvytflsstlksleekdhih 476
           ||||::||||||||||:||||  ::||  |::|:|::||||||::| ::::|   : :    :
Qy     369 CVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSRKLA 427

Db     477 rvldkitdtlihlmakagltlqqqhqrlaqllliilshirhmsnkgmehlysmkcknvvpl 536
           ::|: :||:|: ::||:|:: |||  |||  ||::|||:||  |||||||| :||||||||:
Qy     428 HLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPV 487

Db     537 ydlllemldahrlha-ptsrggasveetdqshlatagstsshs 578
           |||||||||:|| |::   :|  |:    :::|  :  || :::|
Qy     488 YDLLLEMLNAHVLRGCKSSITGSECSPAEDSK-SKEGSQNPQS 529
```

```
RESULT     3
ID   ESTR_MOUSE      STANDARD;      PRT;   599 AA.
AC   P19785;
DT   01-FEB-1991 (REL. 17, CREATED)
DT   01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
DT   01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE   ESTROGEN RECEPTOR (ER).
GN   ESR OR ESTR.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=UTERUS;
RX   MEDLINE; 91042558.
RA   WHITE R., LEES J.A., NEEDHAM M., HAM J., PARKER M.;
RL   MOL. ENDOCRINOL. 1:735-744(1987).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- IN THE ABSENCE OF LIGAND, STEROID HORMONE RECEPTORS ARE THOUGHT
CC       TO BE WEAKLY ASSOCIATED WITH NUCLEAR COMPONENTS; HORMONE BINDING
CC       GREATLY INCREASES RECEPTOR AFFINITY. THE HORMONE-RECEPTOR
CC       COMPLEX APPEARS TO RECOGNIZE DISCRETE DNA SEQUENCES UPSTREAM OF
CC       TRANSCRIPTIONAL START SITES.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; M38651; G193180; -.
DR   PIR; A40061; A40061.
DR   HSSP; P06536; 1GDC.
DR   PROSITE; PS00031; STEROID_FINGER.
DR   TRANSFAC; T00259; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
FT   DOMAIN        1    188       MODULATING.
FT   DNA_BIND    189    254       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     189    209       C4-TYPE.
FT   ZN_FING     225    249       C4-TYPE.
FT   DOMAIN      255    599       STEROID-BINDING.
FT   DOMAIN       64     72       POLY-ALA.
SQ   SEQUENCE   599 AA;  66955 MW;  FED9EAFB CRC32;

  Query Match            42.8%;  Score 1669;  DB 3;  Length 599;
  Best Local Similarity  51.7%;  Pred. No. 0.00e+00;
  Matches  246;  Conservative  106;  Mismatches 104;  Indels  20;  Gaps  14;

Db    108 spsplmllhpppqlspflhphgqqvpyylenepsay-avrdtgppafyrsnsdnrrqngr 166
           : || :|    | :||| |   | |      | | : |:      |             :: :|: :|
Qy     73 TTSPNVLWPTPGHLSP-LVVHRQLSHLYAEPQKSPWCEARSLEHTLPVNRETLKRKVSG- 130

Db    167 erlsssnekgnmimesaketrycavcndyasgyhygvwscegckaffkrsiqghndymcp 226
           :|  :|    |      | ::::||||:|||||||||||||||||||||||||||||:||
Qy    131 NR-CASPVTGP---GSKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICP 186

Db    227 atnqctidknrrkscqacrlrkcyevgmmkggirkdrrggrmlkhkrqrddlegrnemga 286
          ||||||||||||||||||||||||||||||:| | |::| | |:::  |||   : :   |
Qy    187 ATNQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVR--RQRSA-DEQLH-CA 242

Db    287 sgdmraanlwpsplvikhtkknspalsltadqmvsalldaeppmiyseydpsrpfseasm 346
           :  :|::    :| | :    :  |||  ::|:| :||:||||  :     || ||:||||
Qy    243 GKAKRSGG--HAPRV-RELLLD--ALS--PEQLVLTLLEAEPPHVLIS-RPSAPFTEASM 294

Db    347 mglltnladrelvhminwakrvpgfgdlnlhdqvhllecawleilmiglvwrsmehpgkl 406
          |  ||:|||:|:|||||.|||..|||  |:| ||| |||  | | || :| |||  |||||
Qy    295 MMSLTKLADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKL 354

Db    407 lfapnllldrnqgkcvegmveifdmllatssrfrmmnlqgeefvclksiillnsgvytfl 466
          :|||:|:|||::|||||||:·||||||||||·||||  ··||   |··|··||||||··| ::
Qy    355 IFAPDLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLV 414

Db    467 sstlksleekdhihrvldkitdtlihlmakagltlqqqhrrlaqlllilshirhmsnkgm 526
           .:|  . :     . ..|. .||.|. .:||:|:: |||  ||| ||::|||:|| | |:|
Qy    415 TATQDA-DSSRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGM 473

Db    527 ehlysmkcknvvplydlllemldahrlhapasrmgvppeepsqtqlattsstsahs 582
          ||| ||||||||:|||||||||:|| |::  | :  :  |::   :  :| :::|
Qy    474 EHLLNMKCKNVVPVYDLLLEMLNAHVLRGCKSSITGSECSPAEDSKSKEGSQNPQS 529


RESULT     4
ID   ESTR_RAT        STANDARD;      PRT;   600 AA.
AC   P06211;
DT   01-JAN-1988 (REL. 06, CREATED)
DT   01-JAN-1988 (REL. 06, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   ESTROGEN RECEPTOR (ER).
GN   ESR OR ESTR.
OS   RATTUS NORVEGICUS (RAT).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WISTAR;
RA   MURAMATSU M.;
RL   SUBMITTED (MAR-1987) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 87174780.
RA   KOIKE S., SAKAI M.;
RL   NUCLEIC ACIDS RES. 15:2499-2513(1987).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=UTERUS;
RA   MAGGI A.M.A.;
RL   SUBMITTED (JUN-1991) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- IN THE ABSENCE OF LIGAND, STEROID HORMONE RECEPTORS ARE THOUGHT
CC       TO BE WEAKLY ASSOCIATED WITH NUCLEAR COMPONENTS; HORMONE BINDING
CC       GREATLY INCREASES RECEPTOR AFFINITY. THE HORMONE-RECEPTOR
CC       COMPLEX APPEARS TO RECOGNIZE DISCRETE DNA SEQUENCES UPSTREAM OF
CC       TRANSCRIPTIONAL START SITES.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; Y00102; G56111; -.
DR   EMBL; X61098; G56121; -.
DR   PIR; S07379; S07379.
DR   HSSP; P06536; 1GDC.
DR   PROSITE; PS00031; STEROID_FINGER.
DR   TRANSFAC; T00258; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
FT   DOMAIN        1    189       MODULATING.
FT   DNA_BIND    190    255       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     190    210       C4-TYPE.
FT   ZN_FING     226    250       C4-TYPE.
FT   DOMAIN      256    600       STEROID-BINDING.
FT   DOMAIN       64     71       POLY-ALA.
FT   CONFLICT    488    488       N -> T (IN REF. 3).
SQ   SEQUENCE   600 AA;  67030 MW;  400C59B2 CRC32;

  Query Match            42.6%;  Score 1663;  DB 3;  Length 600;
  Best Local Similarity  51.1%;  Pred. No. 0.00e+00;
  Matches  243;  Conservative  110;  Mismatches 103;  Indels  20;  Gaps  14;

Db    109 spsplmllhppphvspflhphghqvpyylenepsay-avrdtgppafyrsnsdnrrqngr 167
           : || :|    | |:|| |   | :      | | : |:      |             :: :|: :|
Qy      73 TISPNVLWPTPGHLSP-LVVHRQLSHLYAEPQKSPWCEARSLEHTLPVNRETLKRKVSG- 130

Db     168 erlsssekqnmimesakctrycavcndyasqyhygvwscegckaffkrsiqghndymcp 227
            :|  :|:   |        | ::::||||:||||||||||||||||||||||||||||:||
Qy     131 NR-CASPVTGP---GSKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICP 186

Db     228 atnqctidknrrkscqacrlrkcyevgmmkggirkdrrggrmlkhkrqrddlegrnegt 287
           ||||||||||||||||||||||||||||:| | |::| | |:::: |  |.        |
Qy     187 ATNQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQRSADE-QLHCA-GK 244

Db     288 sdmraanlwpsplmikhtkknspalsltadqmvsalldaepplyseydpsrpfseasm 347
           :  :|::      :| | :      :   |||  ::|:| :||:|||| :     || || ||||
Qy     245 A--KRSGG--HAPRV-RELLLD--ALS--PEQLVLTLLEAEPPHVLIS-RPSAPFTEASM 294

Db     348 mqlltnladrelvhminwakrvpgfqdlnlhdqvhllecawleilmiglvwrsmehpgkl 407
           |  ||:|||:||||||:|||::|||  :|:| |||:|||   |:|.||.||.|||..||||
Qy     295 MMSLTKLADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKL 354

Db     408 lfapnllldrnqgkcvegmveifdmllatssrfrmmnlqgeefvclksiillnsgvytfl 467
           :|||:|:|||::||||||::||||||||||:||||  ::||   |::|:|::||||||::|  ::
Qy     355 IFAPDLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLV 414

Db     468 sstlksleekdhihrvldkindtlihlmakagltlqqqhrrlaqllliilshirhmsnkgm 527
           ::|   :  :      :  ::|:  : |:|: ::||:|:: |||   ||| ||::|||:|| |||||
Qy     415 TATQDA-DSSRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGM 473

Db     528 ehlysmkcknvvplydlllemldahrlhapasrmgvppeepsqsqltttsstsahs 583
           ||| ||||:||||:|||||||||:|| |::  | :  :   |::    :  :| :::|
Qy     474 EHLLNMKCKNVVPVYDLLLEMLNAHVLRGCKSSITGSECSPAEDSKSKEGSQNPQS 529
```

```
RESULT     5
ID   ESTR_ORYLA     STANDARD;      PRT;   575 AA.
AC   P50241;
DT   01-OCT-1996 (REL. 34, CREATED)
DT   01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   ESTROGEN RECEPTOR (ER).
GN   ESR.
OS   ORYZIAS LATIPES (MEDAKA FISH).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; PISCES; GNATHOSTOMATA;
OC   OSTEICHTHYES; ACTINOPTERYGII; ATHERINIFORMES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RA   OKADA H.;
RL   SUBMITTED (APR-1994) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; D28954; G602326; .
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
FT   DOMAIN        1    140       MODULATING.
FT   DNA_BIND    141    206       C4 TYPE ZINC FINGERS (TWO).
FT   ZN_FING     141    161       C4 TYPE.
FT   ZN_FING     177    201       C4 TYPE.
FT   DOMAIN      207    575       STEROID-BINDING.
FT   DOMAIN      254    262       POLY-GLY.
FT   DOMAIN      275    278       POLY-LEU.
FT   DOMAIN      466    470       POLY-LEU.
FT   DOMAIN      531    534       POLY-GLY.
SQ   SEQUENCE   575 AA;  62987 MW;  D9FE40D0 CRC32.

  Query Match             40.8%;  Score 1591;  DB 3;  Length 575;
  Best Local Similarity   52.8%;  Pred. No. 0.00e+00;
  Matches  252;  Conservative   93;  Mismatches 108;  Indels  24;  Gaps  17;
```

```
Db      38 sstqyysapletnqppse-qslqs-lqsqptsplvfvpssprlspfm-hpp-shhylett 93
           ::  :||:::     : ||:   :|::   |     ||| |: |:: :|||:: |    || | |
Qy      44 PAMTFYSPAVMNYSIPSNVTNLEGGPGRQTTSPNVLWPTPGHLSPLVVHRQLSHLYAEPQ 103

Db      94 stpvyrsshqgasredqcgsredtcslqelgagaagfemakdtrfcavcsdyasgyhy 153
           .|      .      .       .      | |.  |..  .|     .  .|::|||||||||||||||
Qy     104 KSPWCEARSLEHTLPVNRETLKRKVS-GNRCASPVTGPGSK-RDAHFCAVCSDYASGYHY 161

Db     154 gvwscegckaffkrsiqghndymcpatnqctidrnrrkscqacrlrkcyevgmmkggvrk 213
           ||||||||||||||||||||||:||||||||||:|||||||||||||||||||||||:| | |:
Qy     162 GVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRK 221

Db     214 dririlrrdkrrtgvgdqdkvvkgqehktvhydgrkrsstgggggggrlsvtsippeq 273
           .|    |   ||        ..     ..|.   .|   |: :  |  ||  :        |  :  ::|||
Qy     222 ERCGY-RLV-RR------QR--SADEQ--LHCAGKAKRS-GGHAPRVRELL-LDALSPEQ 267

Db     274 vllllqgaepp-ilcsrqklsrpytevtmmtlltsmadkelvhmiawakklpgflqlslh 332
           ::|  |   ||||  :|  ||     :  |:||  :||   ||  :||||||||||:||||:||||::|||
Qy     268 LVLTLLEAEPPHVLISRPS-A-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLF 325

Db     333 dqvlllesswlevlmiglivwrsihcpgklifaqdlildrneqdcvegmteifdmllatas 392
           |||  ||||  |:||||:||:||||    |||||||  ||:|||:||  ||||:  ||||||||||:|
Qy     326 DQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTS 385

Db     393 rfrvlklkpeefvclkaiillnsgafsfctgtmeplhnsaavqsmldtitdalihyisqs 452
           ||| |||     |..|.||.||||.  ...  |.|  ..    :|     :    :|::::|||:    |:  |
Qy     386 RFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSRKLAHLLNAVTDALVWVIAKS 444

Db     453 gylaqeqarrqaqlllllshirhmsnkgmehlysmkcknkvplydlllemldahrlh 509
           |   :|:|:  |  |  ||:||||:||  |||||||  :|||||  ||:|||||||||:||  |:
Qy     445 GISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNAHVLR 501
```

```
RESULT     6
ID   ESTR_SALIR     STANDARD;      PRT;   574 AA.
AC   P16058;
DT   01-APR-1990 (REL. 14, CREATED)
DT   01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   ESTROGEN RECEPTOR (ER).
GN   ESR.
OS   SALMO IRIDEUS (RAINBOW TROUT).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; PISCES; GNATHOSTOMATA;
OC   OSTEICHTHYES; ACTINOPTERYGII; SALMONIFORMES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 91006824.
RA   PAKDEL F., LE GAC F., LE GOFF P., VALOTAIRE Y.;
RL   MOL. CELL. ENDOCRINOL. 71:195-204(1990).
RN   [2]
RP   SEQUENCE OF 150-574 FROM N.A.
RC   TISSUE=LIVER;
RX   MEDLINE; 89127284.
RA   PAKDEL F., LE GUELLEC C., VAILLANT C., LE ROUX M.G., VALOTAIRE Y.;
RL   MOL. ENDOCRINOL. 3:44-51(1989).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- IN THE ABSENCE OF LIGAND, STEROID HORMONE RECEPTORS ARE THOUGHT
CC       TO BE WEAKLY ASSOCIATED WITH NUCLEAR COMPONENTS; HORMONE BINDING
CC       GREATLY INCREASES RECEPTOR AFFINITY. THE HORMONE-RECEPTOR
CC       COMPLEX APPEARS TO RECOGNIZE DISCRETE DNA SEQUENCES UPSTREAM OF
CC       TRANSCRIPTIONAL START SITES.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; M31559; G213784; -.
DR   PIR; A37197; A37197.
DR   PIR; A40070; A40070.
DR   HSSP; P06536; 1GDC.
DR   PROSITE; PS00031; STEROID_FINGER.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
FT   DOMAIN        1    146       MODULATING (POTENTIAL).
FT   DNA_BIND    147    212       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     147    167       C4-TYPE.
FT   ZN_FING     183    207       C4-TYPE.
FT   DOMAIN      213    574       STEROID-BINDING (POTENTIAL).
SQ   SEQUENCE   574 AA;  62872 MW;  B391ADC2 CRC32;

  Query Match              38.8%;  Score 1516;  DB 3;  Length 574;
  Best Local Similarity  56.3%;  Pred. No. 0.00e+00;
  Matches  225;  Conservative   78;  Mismatches   75;  Indels   22;  Gaps   17;

Db    143 etrycavcsdfasgyhygvwscegckaffkrsiqghndymcpatnqctmdrnrrkscqac 202
          ::::|||||||:|||||||||||||||||||||||||||:||||||||:|:||||||||||
Qy    145 DAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQAC 204

Db    203 rlrkcyevgmmkggirkdrgg-rvlrkdkrygqpagdrekpytvtwstggrpdggrnsss 261
          |||||||||||| | |::| | |::|:: |   :| | |     : : : |:  ||:  ::
Qy    205 RLRKCYEVGMVKCGSRRERCGYRLVRRQ-R---SA-D-EQLHCA-GKAKRS--GGH--AP 253

Db    262 slnqggwrqprltmppeqvlfllqgqtp-alqsrqkvarpytevtmmtlltsmadkelv 320
            :        : :: | : || ||::: |  | ||    | |:|| :|| || :||||||
Qy    254 RVREL--LLDA-LS-PEQLVLTLLEAEPPHVLISRPS-A-PFTEASMMMSLTKLADKELV 307

Db    321 hmiawakkvpgfqelslhdqvqllesswlevlmigllwrslhcpgklifaqdlildrseg 380
          |||:||||:||| |||| |||:|||| |:||||:||:||||  ||||||| ||:||| ||
Qy    308 HMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEG 367

Db    381 dcvegmaeifdmllatvsrfrmlklkpeefvclkaiillnpgafsfcsnsveslhnssav 440
           ||||: |||||||| |||  |||   |::|:||:|||:: ::: : : :: :|  :
Qy    368 KCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSRKL 426

Db    441 esmldnitdalihhishsgasvqqqprrqaqlllllshirhmsnkgmehlysikcknkvp 500
            :|: :||||:  !: || | |||: | | ||:||||:|| ||||||| ::|||| ||
Qy    427 AHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVP 486

Db    501 lydlllemldghrlqapgkvaqageqtegpstttttstgs 540
          :|||||||::| |::  | : :| :  :|:  : :  ||
Qy    487 VYDLLLEMLNAHVLRGC-KSSITGSEC-SPAEDSKSKEGS 524


RESULT    7
ID   ESTR_SALSA     STANDARD;      PRT;   535 AA.
AC   P50242;
DT   01-OCT-1996 (REL. 34, CREATED)
DT   01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   ESTROGEN RECEPTOR (ER) (FRAGMENT).
GN   ESR.
OS   SALMO SALAR (ATLANTIC SALMON).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; PISCES; GNATHOSTOMATA;
OC   OSTEICHTHYES; ACTINOPTERYGII; SALMONIFORMES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RA   ROGERS S.A., LLEWELLYN L., RAMSURN V.R., SWEENEY G.E., WIGHAM T.;
RL   SUBMITTED (JUL-1995) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; X89959; G929902; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
FT   NON_TER       1      1
FT   DOMAIN       <1    104       MODULATING (POTENTIAL).
FT   DNA_BIND    105    170       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     105    125       C4-TYPE.
FT   ZN_FING     141    165       C4-TYPE.
FT   DOMAIN      171    535       STEROID-BINDING (POTENTIAL).
SQ   SEQUENCE   535 AA;  58900 MW;  0F71CD0C CRC32;

  Query Match              38.5%;  Score 1503;  DB 3;  Length 535;
  Best Local Similarity  55.7%;  Pred. No. 0.00e+00;
  Matches  224;  Conservative   75;  Mismatches   80;  Indels   23;  Gaps   15;

Db    101 etrycavcsdfasgyhygfwscegckaffkrsiqghndymcpatnqctmdrnrrkscqac 160
          ::::|||||||:|||||| |||||||||||||||||||||||||:||||||||:|:||||||||
Qy    145 DAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQAC 204

Db    161 rlrkcyevgmmkggirkdrgg-rvlrkdkrygqpagdrekpygdlehrtappqdggrnss 219
          |||||||||||| | |..| | |..|.. |   .| . .   .      : |  : ||: :
Qy    205 RLRKCYEVGMVKCGSRRERCGYRLVRRQ-R---SADE-QLHCAG--K-AK-RSGGHAPR 254

Db    220 ssslsggggwgprltmppeqvlfllqgaeppalcsrqkvarpytevtmmtlltsmadke 278
             |        : :: | : || ||::  |  | ||   | |:|| :||  || :||||
Qy    255 VRELL----L-DA-LS-PEQLVLTLLEAEPPHVLISRPS-A-PFTEASMMMSLTKLADKE 305

Db    280 lvhmiawakkvpgfqelslhdqvqllesswlevlmigllwrslhcpgklifaqdlildrs 339
          |||||:||||:||| |||| |||:|||| |:||||:||:||||  ||||||| ||:|||
Qy    306 LVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRD 365

Db    340 egdcvegmaeifdmllatvsrfrmlklkpeefvclkaiillnsgafsfcsnsveslhnss 399
          || ||||: |||||||| |||| |||    |::|:||:|||||: ::: : : :: :'
Qy    366 EGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSR 424

Db    400 avesmldnitdalihhishsgasvqqqprrqvqlllllshirhmsnkgmehlysikcknk 459
            : :|: :||||:  |: || | |||: |   ||:||||:|| |||||.||| ::||||
Qy    425 KLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNV 484

Db    460 vplydlllemldghrlqspgkvaqageqtegpstttttstgs 501
          ||:||||||||::| |::  | : :| :  :|:  : :  ||
Qy    485 VPVYDLLLEMLNAHVLRGC-KSSITGSEC-SPAEDSKSKEGS 524


RESULT    8
ID   ESTR_OREAU     STANDARD;      PRT;   314 AA.
AC   P50240;
DT   01-OCT-1996 (REL. 34, CREATED)
DT   01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   ESTROGEN RECEPTOR (ER) (FRAGMENT).
GN   ESR.
OS   OREOCHROMIS AUREUS.
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; PISCES; GNATHOSTOMATA;
OC   OSTEICHTHYES; ACTINOPTERYGII; PERCIFORMES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BLOOD;
RX   MEDLINE; 96272302.
RA   TAN N., LAM T., DING J.;
RL   DNA SEQ. 5:359-370(1995).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; Z46665; E212276; -.
DR   EMBL; Z46666; E212276; JOINED.
DR   EMBL; Z46667; E212276; JOINED.
DR   EMBL; Z46668; E212276; JOINED.
DR   EMBL; Z46669; E212276; JOINED.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
```

```
FT   NON_TER       1      1
FT   DOMAIN       <1    314       STEROID BINDING (POTENTIAL).
SQ   SEQUENCE   314 AA;  34872 MW;  0B18C899 CRC32;

  Query Match            27.6%;  Score 1076;  DB 3;  Length 314;
  Best Local Similarity  60.3%;  Pred. No. 3.03e-224;
  Matches  146;  Conservative   45;  Mismatches  46;  Indels   5;  Gaps   5;

Db       1 vllllqqaeppilcsrqkmsrpytevtimtlltsmadkelvhmitwakklpgflqlslhd 60
           || ||::  | :| ||    : |:|| ::|   || :||||||||||:||||:|||::||| |
Qy     269 VLTLLEAEPPHVLISRPS-A-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 326

Db      61 qvlllesswlevlmiqliwrsiqcpgklifeedlildrneqtcveqmaeifdmllatv-r 119
           || |||| |:||||:||:||||: ||||||  ||:|||:|| ||||: |||||||||  |
Qy     327 QVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSR 386

Db     120 frvlklkpeefvclkaiillnsgafsfctgtmeplhdsaavqhmldtitdalifhishfg 179
           || |||    |::|:||:|||||: ::: |:| ::    |   : |:|:::||||:  |:  |
Qy     387 FRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSRKLAHLLNAVTDALVWVIAKSG 445

Db     180 csaqqqsrrqaqllllshirhmsnkgmehlysmkckgkvplydlllemldahrihrpvk 239
            |:|||| | |  ||:||||:|| |||||||| :||||| ||:|||||||||:|| : |  |
Qy     446 ISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNAH-VLRGCK 504

Db     240 ps 241
           :|
Qy     505 SS 506


RESULT      9
ID   ERR2_HUMAN      STANDARD;      PRT;   433 AA.
AC   P11475;
DT   01-OCT-1989 (REL. 12, CREATED)
DT   01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)
DT   01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE   STEROID HORMONE RECEPTOR ERR2
GN   ERR2.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=HEART;
RX   MEDLINE; 88122546.
RA   GIGUERE V., YANG N., SEGUI P., EVANS R.M.;
RL   NATURE 331:91-94(1988).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; X51417; G36611; -.
DR   PIR; B29345; B29345.
DR   HSSP; P06536; 1GDC.
DR   PROSITE; PS00031; STEROID_FINGER.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER.
FT   DNA_BIND    103    168       C4-TYPE ZINC FINGERS (TWO)
FT   ZN_FING     103    123       C4-TYPE.
FT   ZN_FING     139    163       C4-TYPE
SQ   SEQUENCE   433 AA;  48287 MW;  673BF88F CRC32;

  Query Match            20.7%;  Score 807;  DB 3;  Length 433;
  Best Local Similarity  39.5%;  Pred. No. 4.26e-158;
  Matches  140;  Conservative   84;  Mismatches 108;  Indels  22;  Gaps  14;

Db     101 rlclvcgdiasgyhygvasceackaffkrtiqgnieyscpatnecitkrrrkscqacrf 160
           ::| ||:| |||||||| |||:||||||||:|||: :| |||||:| | | |||||||||:
Qy     147 HFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRL 206

Db     161 mkclkvgmlkegvrldrv--r-ggrqky-krrldsenspylslqisppakkpltkivs- 214
            ||  |||:| | | :|    |   ||:      :|       :  |   :|  :  |    :|
Qy     207 RKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPE 266

Db     215 ylllvaepdklyamppddvpegdikalttlcdladrelvvllswakhipgfsnltlgd 271
```

```
           | | : |::     : :    |  :     : :|  |||:||| :||||| |||| :|:: |
Qy     267 QLVLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 326

Db     272 qmsllqsawmeililgivyrslpyddklayaedyimdechsrlv-gllelyrailqlvrr 330
           |. ||.| |||.|..|.. ||.    || .| | ::| : :: | |:||::  :|     |
Qy     327 QVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGIIEIFDMLLATTSR 386

Db     331 ykklkvekeefvmlkalalansdsmy--i---enleavqklqdllh---ealq-dyels- 380
           :: ||::  |:: :||: | || |||   :     :: :: :||  ||:    :||       |
Qy     387 FRELKLQHKEYLCVKAMILLNS-SMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKSG 445

Db     381 -qrheeprraqklltlpllrqtaakavqhfysvklqgkvpmhklflemleakv 433
               ....  |  ..||. |. .|... |...|. ..|       ||:  |:||||:| |
Qy     446 ISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNAHV 499


RESULT     10
ID   ERR1_HUMAN      STANDARD;      PRT;   521 AA.
AC   P11474;
DT   01-OCT-1989 (REL. 12, CREATED)
DT   01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)
DT   01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE   STEROID HORMONE RECEPTOR ERR1.
GN   ERR1.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=KIDNEY;
RX   MEDLINE; 88122546.
RA   GIGUERE V., YANG N., SEGUI P., EVANS R.M.;
RL   NATURE 331:91-94(1988).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS
DR   EMBL; X51416; G36609; -.
DR   PIR; A29345; A29345.
DR   HSSP; P06536; 1GDC.
DR   PROSITE; PS00031; STEROID_FINGER.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER.
FT   DNA_BIND    176    241       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     176    196       C4-TYPE.
FT   ZN_FING     212    236       C4-TYPE.
SQ   SEQUENCE   521 AA;  55742 MW;  E2D723EA CRC32;

  Query Match            20.5%;  Score 800;  DB 3;  Length 521;
  Best Local Similarity  39.9%;  Pred. No. 2.06e-156;
  Matches  141;  Conservative   74;  Mismatches 128;  Indels  10;  Gaps   7;

Db     174 rlclvcgdvasgyhygvasceackaffkrtiqgsieyscpasneceitkrrrkacqacrf 233
           ::| ||:| |||||||| |||:||||||||:|||  :| |||:|:| | | |||:|||||:
Qy     147 HFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRL 206

Db     234 tkclrvgmlkegvrldrvrggrqkykrrpevdplpfpg-p-fpagplavaggprktaapv 291
            ||  |||:| | | :|  | |    ::|:  : |  :| :  ::|  :           | :
Qy     207 RKCYEVGMVKCGSRRERC-GYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSP 265

Db     292 nalvshllvvepeklyampdpagpdghlpavatlcdlfdreivvtiswakslpgfsslsl 351
           : ||  ||  ||  :  :: |::|     : : :|  | |:|:| |||||:|||: |||
Qy     266 ELVLTLLEAEPPHVL-ISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSL 324

Db     352 sdqmsvlqsvwmevlvlgvaqrslplqdelafaedlvld-eegaraaglqelgaallqlv 410
            ||. .|.| |||||..|: ||:        | || ||||| :||      |: |:    ||
Qy     325 FDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGIIEIFDMLLATT 384

Db     411 rrlqalrlereeyvllkalalansdsvhiedeprlwsscekllh--eal---leyeagra 465
            |.. |.|.. ||. .||: | ||      :      :    |  ||| |  :|:    |     :::
Qy     385 SRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKS 444

Db     466 gpgggaerrragrllltlpllrqtaqkvlahfyqvklegkvpmhklflemlea 518
           | ::  :  | : ||: |: :|::: | : |:  :|     ||:  |:||||:|
Qy    445 GISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNA 497

RESULT   11
ID   PRGR_MOUSE     STANDARD;      PRT;   923 AA.
AC   Q00175;
DT   01-DEC-1992 (REL. 24, CREATED)
DT   01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   PROGESTERONE RECEPTOR (PR).
GN   PGR OR PR.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 91299759.
RA   SCHOTT D.R., SHYAMALA G., SCHNEIDER W., PARRY G.;
RL   BIOCHEMISTRY 30:7014-7020(1991).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; M68915; G200472; -.
DR   PIR; A39536; A39536.
DR   HSSP; P06536; 1GDC.
DR   PROSITE; PS00031; STEROID_FINGER.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
FT   DOMAIN        1    556       MODULATING, PRO-RICH.
FT   DNA_BIND    557    622       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     557    577       C4-TYPE.
FT   ZN_FING     593    617       C4-TYPE.
FT   DOMAIN      671    923       STEROID-BINDING.
FT   DOMAIN      184    188       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
SQ   SEQUENCE   923 AA;  99073 MW;  EEDFF41A CRC32;

  Query Match             17.4%;  Score 681;  DB 7;  Length 923;
  Best Local Similarity   32.9%;  Pred. No. 1.38e-127;
  Matches  122;  Conservative  112;  Mismatches 128;  Indels  27;  Gaps  27;

Db    491 aapgtaaapaiygplg-lnglpqlg-y qaavlkdslpqvypqylnylrpdseasqspqy 547
           ::||  ::  : : |: | : ||  :   :|  |  :||  |     |:        ||
Qy     81 PTPGHLSPLVVHRQLSHLYAEPQKSPWCEAPSLEHTLP-VNRETLK-RKVSGNRCASPVT 138

Db    548 gfdslpq-kiclicgdeasgchygvltcgsckvffkrameggqhnylcagrndcivdkirr 606
           |  |  :  :| :|:| ||| |||| :| :|| |||':: :|::.|:.|.. |:| .|| ||
Qy    139 GFGSKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRR 198

Db    607 kncpacrlrkccqagmvl-ggrkfk kfnkvrvmrtldgvalpqsvglpnesqalsqrit 664
           |:| |||||||| . ||| |.|.        || |  |   | .  |  :  |  : :
Qy    199 KSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQRSADE-QL-HCAGKAKRSGGHAPRVR 256

Db    665 fspnqniqlvp-plinllmsiepdviyaghdntkpdtsssllts lnqlgerqllsvvkws 723
                :  |   ::  |:  ||  .          | |  |  ||  |:  |   |:
Qy    257 -ELLIDA-ISPEQLVLTLLEAEPPHVLISRPSA-PFTEASMMMSLTKLADKELVHMISWA 313

Db    724 kslpgfrnlhiddqitliqyswmslmvfglgwrsykhvsgqmlyfapdlilne-q-r-mk 780
           | :||| :| : ||: '..  || ... || |||  | .| .. |||||.|| .. : :
Qy    314 KKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVE 371

Db    781 e-lsfyslcltmwqipqefvklqvtheeflcmkvllllnt-i-pleglrsqsqfeemrss 837
           | :: : | :       |  | : | |:||.| ..:||. . ||       :::    . .
Qy    372 GILEIFDM-L-LATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAH 428

Db    838 yirelikaig-lrqkgvvptsqrfyqltklldslhdlvkq 876
            :  :  |:  :  |: ::: |:  :|::||  |   .::
Qy    429 LLNAVTDALVWVIAKSGISSQQQSMRLANLL-MLLSHVRH 467
```

```
RESULT   12
ID   PRGR_HUMAN     STANDARD;      PRT;   933 AA.
AC   P06401;
DT   01-JAN-1988 (REL. 06, CREATED)
DT   01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   PROGESTERONE RECEPTOR (PR).
GN   PGR.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 90228361.
RA   KASTNER P., KRUST A., TURCOTTE B., STROPP U., TORA L., GRONEMEYER H.,
RA   CHAMBON P.;
RL   EMBO J. 9:1603-1614(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 87184565.
RA   MISRAHI M., ATGER M., D'AURIOL L., LOOSFELT H., MERIEL C.,
RA   FRIDLANSKY F., GUIOCHON-MANTEL A., GALIBERT F., MILGROM E.;
RL   BIOCHEM. BIOPHYS. RES. COMMUN. 143:740-748(1987).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- ALTERNATIVE PRODUCTS: THERE ARE TWO FORMS OF THE PROGESTERONE
CC       RECEPTOR: FORM A (THAT STARTS WITH MET-1) AND FORM B (THAT STARTS
CC       WITH MET-165).
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; X51730; G35652; -.
DR   EMBL; M15716; G189935; -.
DR   PIR; A03245; QRHUP.
DR   PIR; S09971; S09971.
DR   HSSP; P06536; 1GDC.
DR   MIM; 264080; -.
DR   PROSITE; PS00031; STEROID_FINGER.
DR   TRANSFAC; T00696; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING; PHOSPHORYLATION; ALTERNATIVE SPLICING.
FT   DOMAIN        1    566       MODULATING, PRO-RICH.
FT   DNA_BIND    567    632       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     567    587       C4-TYPE.
FT   ZN_FING     603    627       C4-TYPE.
FT   DOMAIN      681    933       STEROID-BINDING.
FT   DOMAIN      183    187       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT   MOD_RES      41     41       PHOSPHORYLATION (BY CK).
FT   MOD_RES     227    227       PHOSPHORYLATION (BY CK).
FT   MOD_RES     232    232       PHOSPHORYLATION (BY CK).
FT   MOD_RES     552    552       PHOSPHORYLATION (BY CK).
FT   MOD_RES     794    794       PHOSPHORYLATION (BY CK).
FT   VARSPLIC      1    164       MISSING (IN FORM B).
FT   CONFLICT    229    229       E -> G (IN REF. 2).
FT   CONFLICT    259    259       G -> V (IN REF. 2).
FT   CONFLICT    344    344       T -> S (IN REF. 2).
FT   CONFLICT    660    660       L -> V (IN REF. 2).
SQ   SEQUENCE   933 AA;  99042 MW;  0E1124EB CRC32;

  Query Match             17.1%;  Score 666;  DB 7;  Length 933;
  Best Local Similarity   33.4%;  Pred. No. 5.61e-124;
  Matches  110;  Conservative   89;  Mismatches 113;  Indels  17;  Gaps  17;

Db    566 lclicgdeasgchygvltcgsckvffkrameggqhnylcagrndcivdkirrkncpacrlr 625
           .| .|.| ||| |||| .| .|| ||||.  |    | | : |:| :|| |||:| |||||
Qy    148 FCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLR 207

Db    626 kccqagmvl ggrkfk-kfnkvrvvraldavalpqp-gvpnesqalsqrftfspgqdiq. 683
           || : |||  |:|: :  :  ||  |: |    | :  | :: | : :        : |
Qy     208 KCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSP 265

Db     684 ip-plinllmsiepdvlyaghdntkpdtsssslltslnqlqerqllsvvkwskslpgfrnl 742
              |  ':   |:   ||    :    :: ::  | |  :|:: ||   |::::|: :: |:| :||| :|
Qy     264 SPEQLVLTLLEAEPPHVLISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVEL 322

Db     743 hiddqitliqyswmslmvfglgwrsykhvsgqmlyfapdlilneqrmk-essfyslcltm 801
             :  ||: |::   || ::: || |||    | :| :: |||||:|: :   |     ::  :   :
Qy     323 SLFDQVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDML 380

Db     802 wqlpqefvklqvsqeeflcmkvllllntipleglrsqtqfeemrssyirelikaig- 857
                    |    |  :  : |:||:| ::|||: ::   | : ||   :   |   :  :    :  |:
Qy     381 LATTSRFRELKLQHKEYLCVKAMILLNS-SMYPLVTATQDADSSRKLAHLLNAVTDALVW 439

Db     858 lrqkqvvsssqrfyqltklldnlhdlvkq 886
             :  |: :|| |:    :|::||   |    |::
Qy     440 VIAKSGISSQQQSMRLANLL-MLLSHVRH 467


RESULT   13
ID   PRGR_RABIT     STANDARD;      PRT;   930 AA
AC   P06186;
DT   01-JAN-1988 (REL. 06, CREATED)
DT   01-JAN-1988 (REL. 06, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   PROGESTERONE RECEPTOR (PR).
GN   PGR.
OS   ORYCTOLAGUS CUNICULUS (RABBIT).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; LAGOMORPHA.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 87067449.
RA   LOOSFELT H., ATGER M., MISRAHI M., GUIOCHON-MANTEL A., MERIEL C.,
RA   LOGEAT F., BENAROUS R., MILGROM E.;
RL   PROC. NATL. ACAD. SCI. U.S.A. 83:9045-9049(1986).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; M14547; G165632; -.
DR   PIR; A25923; A25923.
DR   HSSP; P06536; 1GDC.
DR   PROSITE; PS00031; STEROID_FINGER.
DR   TRANSFAC; T00697; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC FINGER; STEROID-BINDING.
FT   DOMAIN         1     565       MODULATING, PRO-RICH.
FT   DNA_BIND     566     633       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING      568     588       C4-TYPE.
FT   ZN_FING      604     628       C4-TYPE.
FT   DOMAIN       678     930       STEROID-BINDING.
FT   DOMAIN       184     188       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
SQ   SEQUENCE   930 AA;  98666 MW;  EF133E86 CRC32;

  Query Match             17.0%;  Score 662;  DB 7;  Length 930;
  Best Local Similarity   32.5%;  Pred. No. 5.13e-123;
  Matches  106;  Conservative   51;  Mismatches 114;  Indels  15;  Gaps  15;

Db     567 iclicgdeasgchygvltcgsckvffkrameqqhnylcagrndcivdkirrkncpacrlr 626
            :|  :|:| ||| |||| :|  :|| ||||:::|:::|:|:: |:| :|| |||:| |||||
Qy     148 FCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLR 207

Db     627 kccqagmvl-ggrkfk kfnkvrvmral dvalpqpvgipnesqrltfspsqdlq lip 682
            || : |||  |:|: :  :  ||  |: |    | :  | :: |   :          :: | |
Qy     208 KCYEVGMVKCGSRRERCGYRLVRRQRSADE QLHCAGKAKRSGGHAPRVRELLLDALSP 265

Db     683 -plinllmsiepdviyaghdntkpdtsssllts lnqlgerqllsvvkwskslpgfrnlhi 741
              |:  |:  ||  :  :: :: |   :|:: ||  |::::|: :: |:| :||| :| :
```

```
Qy     266 EQLVLTLLEAEPPHVLISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSL 324

Db     742 ddqitliqyswmslmvfglgwrsykhvsgqmlyfapdlilneqrmk-essfyslcltmwq 800
             ||: |::   || ::: || |||    | :| :: |||||:|: :   |     ::  :   :
Qy     325 FDQVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDMLLA 382

Db     801 ipqefvklqvsqeeflcmkvllllnt-i-pleglrsqsqfeemrssyirelikaiq-lrq 857
                 |   |  :  : |:||:| ::|||: :  ||      :::    :  :   :   :  |:  :
Qy     383 TTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIA 442

Db     858 kgvvsssqrfyqltklldnlhdlvkq 883
            |: :||  |:   :|::||   |    |::
Qy     443 KSGISSQQQSMRLANLL-MLLSHVRH 467

RESULT   14
ID   GCR_HUMAN      STANDARD;      PRT;   777 AA.
AC   P04150; P04151;
DT   01-NOV-1986 (REL. 03, CREATED)
DT   01-NOV-1986 (REL. 03, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   GLUCOCORTICOID RECEPTOR (GR).
GN   GRL.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 86092206.
RA   HOLLENBERG S.M., WEINBERGER C., ONG E.S., CERELLI G., ORO A.,
RA   LEBO R., THOMPSON E.B., ROSENFELD M.G., EVANS R.M.;
RL   NATURE 318:635-641(1985).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   MUNROE D.G., PANG L., TAYLOR G.R., LAU C., PLANTE R.K., ZHOU L.;
RL   SUBMITTED (SEP-1993) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [3]
RP   SEQUENCE OF 1-394 FROM N.A.
RX   MEDLINE; 91224961.
RA   LECLERC S., XIE B.X., ROY R., GOVINDAN M.V.;
RL   J. BIOL. CHEM. 266:8711-8719(1991).
RN   [4]
RP   DOMAINS.
RX   MEDLINE; 86092211.
RA   WEINBERGER C., HOLLENBERG S.M., ROSENFELD M.G., EVANS R.M.;
RL   NATURE 318:670-672(1985).
RN   [5]
RP   VARIANT SER-363.
RX   MEDLINE; 93187003.
RA   KARL M., LAMBERTS S.W.J., DETERA-WADLEIGH S.D., ENCIO I.J.,
RA   STRATAKIS C.A., HURLEY D.M., ACCILI D., CHROUSOS G.P.;
RL   J. CLIN. ENDOCRINOL. METAB. 76:683-689(1993).
RN   [6]
RP   VARIANT VAL-641.
RX   MEDLINE; 91123468.
RA   HURLEY D.M., ACCILI D., STRATAKIS C.A., KARL M., VAMVAKOPOULOS N.,
RA   RORER E., CONSTANTINE K., TAYLOR S.I., CHROUSOS G.P.;
RL   J. CLIN. INVEST. 87:680-686(1991).
RN   [7]
RP   VARIANT ILE-729.
RX   MEDLINE; 93253031.
RA   MALCHOFF D.M., BRUFSKY A., REARDON G., MCDERMOTT P., JAVIER E.C.,
RA   BERGH C.H., ROWE D., MALCHOFF C.D.;
RL   J. CLIN. INVEST. 91:1918-1925(1993).
RN   [8]
RP   VARIANT PHE-753.
RX   MEDLINE; 93302771.
RA   ASHRAF J., THOMPSON E.B.;
RL   MOL. ENDOCRINOL. 7:631-642(1993).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES. BINDS TO THE
CC       GRE TARGET SITE.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS
CC   -!- ALTERNATIVE PRODUCTS: ALPHA AND BETA FORMS OF HGR DIFFER ONLY IN
CC       THEIR CARBOXYL-TERMINAL SEQUENCES. THEY ARE GENERATED BY
CC       ALTERNATIVE SPLICING. ALPHA-HGR IS THE PREDOMINANT PHYSIOLOGICAL
CC       FORM OF HGR. THE IN VITRO TRANSLATION PRODUCT OF HGR BETA DOES NOT
CC       BIND STEROID ITS FUNCTION IS UNCLEAR; HOWEVER, IT IS POSSIBLE THAT
CC       VARIANT RECEPTORS PERFORM TISSUE-SPECIFIC FUNCTIONS.
CC   -!- DISEASE: DEFECTS IN GRL ARE A CAUSE OF FAMILIAL GLUCOCORTICOID
CC       RESISTANCE (CORTISOL RESISTANCE): A HYPERTENSIVE, HYPERANDROGENIC
CC       DISORDER CHARACTERIZED BY INCREASED SERUM CORTISOL CONCENTRATIONS
DR   EMBL; M10901; G183033; -.
DR   EMBL; M11050; G183035; -
DR   EMBL; U01351; G458657; -.
DR   EMBL; X03225; G31680; -.
DR   EMBL; X03348; G31682; -.
DR   EMBL; M69104; G553322; -.
DR   EMBL; M73816; G553323; -.
DR   EMBL; S68378; G239758; -.
DR   PIR; A03246; QRHUGA.
DR   PIR; A03247; QRHUGB.
DR   HSSP; P06536; 1GDC
DR   MIM; 138040; -.
DR   PROSITE; PS00031; STEROID_FINGER.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING; DISEASE MUTATION; POLYMORPHISM;
KW   ALTERNATIVE SPLICING.
FT   DOMAIN      145    278       IMMUNOGENIC.
FT   DNA_BIND    421    486       C4-TYPE ZINC FINGERS (TWO)
FT   ZN_FING     421    441       C4-TYPE.
FT   ZN_FING     457    481       C4-TYPE.
FT   DOMAIN      487    527       HINGE.
FT   DOMAIN      528    777       STEROID-BINDING.
FT   VARIANT     363    363       N -> S.
FT   VARIANT     641    641       D -> V (IN GLUCOCORTICOID RESISTANCE)
FT   VARIANT     729    729       V -> I (IN GLUCOCORTICOID RESISTANCE)
FT   VARIANT     753    753       L -> F (IN GLUCOCORTICOID RESISTANCE)
FT   VARSPLIC    728    777       VVENLLNYCFQTFLDKTMSIEFPEMLAEIITNQIPKYSNGN
FT                                IKKLLFHQK -> ENVMWLKPESTSHTLI (IN BETA
FT                                FORM).
FT   CONFLICT    451    451       G -> GR (IN REF. 2)
SQ   SEQUENCE   777 AA;  85659 MW;  D93D1F01 CRC32;

  Query Match              15.7%;  Score 614;  DB 4;  Length 777;
  Best Local Similarity   33.4%;  Pred. No. 1.61e-111;
  Matches   97;  Conservative   84;  Mismatches   93;  Indels   16;  Gaps   15;

Db     397 smrpdvssppssstattgp-p-pklclvcsdeasgchygvltcgsckvffkraveqqhn 454
           :::  ||:   ::|:  :   |        |  |||| ||: ||||  |  || |||:  :::  :
Qy     123 TIKPKVSGNPCASPVTGPGSKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHND 182

Db     455 ylcagrndciidkirrkncpacryrkclqagmnl-earktk-k-kikgiqqattg-v-sq 509
           |:|:: |:| ||| |||:| ||| ||| : ||      :|:  :   ::   |:::    :
Qy     183 YICPATNQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCA 242

Db     510 etsenpgnkt-ivpatlpq-ltpt-lv-slleviepevlyagydssvpdstwrimttlnm 565
             :  :|  :   |     | : |.|  || .|||    | ||  :             : |  .
Qy     243 GKAKRSGGHAPRVRELLLDALSPEQLVLTLLEAEPPHVLISRPSAPFTEAS--MMMSLIK 300

Db     566 lqgrqviaavkwakaipgfrnlhlddqmtllqyswmflmafalgwrsyrqssanllcfap 625
           |: :...   . ||| |||| .| | ||. ||:  || ::  :| |||  ::: :|: |||
Qy     301 LADKEIVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPG-KLI-FAP 358

Db     626 dliin-eqrmtlpcmydqckhmlyvsselhrlqvsyeeylcmktlllss 674
           ||::: :: : :  :  .      :|  .| .. | .   ||||.|:::||:|
Qy     359 DLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNS 408

RESULT   15
ID   GCR_XENLA      STANDARD;      PRT;   776 AA.
```

```
AC   P49844;
DT   01-OCT-1996 (REL. 34, CREATED)
DT   01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   GLUCOCORTICOID RECEPTOR (GR).
OS   XENOPUS LAEVIS (AFRICAN CLAWED FROG).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AMPHIBIA; ANURA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=EMBRYO;
RX   MEDLINE; 94289478.
RA   GAO X., KALKHOVEN E., PETERSON-MADURO J., VAN DER BURG B.,
RA   DESTREE O.H.J.;
RL   BIOCHIM. BIOPHYS. ACTA 1218:194-198(1994).
RN   [2]
RP   SEQUENCE OF 452-776 FROM N.A.
RC   TISSUE=LIVER;
RA   PICARD D.;
RL   SUBMITTED (FEB-1994) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES. BINDS TO THE
CC       GRE TARGET SITE.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; X72211; G444043; -.
DR   EMBL; X77764; G456374; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
FT   DOMAIN        1    419       MODULATING.
FT   DNA_BIND    420    485       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     420    440       C4-TYPE.
FT   ZN_FING     456    480       C4-TYPE.
FT   DOMAIN      486    776       STEROID-BINDING.
FT   CONFLICT    502    502       S -> P (IN REF. 2).
FT   CONFLICT    504    504       T -> A (IN REF. 2).
FT   CONFLICT    506    506       T -> A (IN REF. 2)
FT   CONFLICT    511    511       P -> N (IN REF. 2).
FT   CONFLICT    519    519       L -> M (IN REF. 2).
FT   CONFLICT    544    544       Y -> F (IN REF. 2).
FT   CONFLICT    551    551       I -> M (IN REF. 2).
FT   CONFLICT    654    654       S -> R (IN REF. 2).
FT   CONFLICT    740    740       L -> M (IN REF. 2).
FT   CONFLICT    768    768       L -> I (IN REF. 2).
SQ   SEQUENCE   776 AA;  84977 MW;  8960496D CRC32;

  Query Match              15.6%;  Score 608;  DB 4;  Length 776;
  Best Local Similarity   34.9%;  Pred. No. 4.36e-110;
  Matches   99;  Conservative   74;  Mismatches   96;  Indels   15;  Gaps   15;

Db     398 irsdaspspstsstgpppklclvcsdeasgchygvltcgsckvffkraveqqhnylca 457
           : ::  :|| |:: |     : :| |||| ||| |||| :| :|| ||| ::: :::|: :
Qy     128 VSGNPCASPVTGPGSK-RDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICP 186

Db     459 grndciidkirrkncpacryrkclqagmnl-earktk-k-kikgiqqsttatar-es-pc 512
           : |:| ||| |||:| ||| ||| : ||      :|: :   ::   |:|:     :  :  :
Qy     187 ATNQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAK 246

Db     513 tsmtrtl-vpasv-aqltpt-li-slleviepevlysgydssipdttrrlmsslnmlqgr 568
            |  ::  |   :   |:|  |: :|||    | || |   ::: :::  :|| | :|: :
Qy     247 RSGGHAPRVRELLLDALSPEQLVLTLLEAEPPHVLISRPSAPFTEASM-MMS-LIKLADK 304

Db     569 qvvsavrwakaipgfrnlhlddqmtllqyswmflmvfalgwrsykqtngsilyfapdlvi 628
           ..|  : ||| |||| .| | ||. ||.  || ... .| |||   .   |   | ||||||:
Qy     305 ELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHP-GK-LIFAPDLVL 362

Db     629 t-edrmhlpfmqercqemlkiagemsslqisydeylcmkvlllm 671
             :: : :  : |  : :|  ::    | :   ||||:| ::|:
Qy     363 DRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILL 406
```

```
Search completed: Thu Dec 18 09:17:23 1997
Job time : 33 secs.
```

```
*********************************************************************
MPSRCH                                                         (TM)
*********************************************************************

          Release 2 1D John F  Collins  Biocomputing Research Unit
          Copyright (c) 1993, 1994  1995  University of Edinburgh, U K
                 Distribution rights by IntelliGenetics, Inc.

MPsrch_pp   protein - protein database search  using Smith-Waterman algorithm

Run on           Thu Dec 18 09 17 49 1997,  MasPar time 20 53 Seconds
                                            743.476 Million cell updates/sec
Tabular output not generated.

Title:           >US-08-826-361A-25
Description:     (1-530) from US08826361A.pep
Perfect Score:   3904
Sequence         1 MDIKNSPSSINSPSSYNCSQ          ECSPAEDSKSKEGSQNPQSQ 530

Scoring table:   PAM 150
                 Gap 11

Searched:        91006 seqs. 28883923 residues

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database:        pir51
                   1:ann1 2:ann2 3:ann3 4:ann4 5:unann1 6:unann2 7:unann3
                   8:unann4 9:unann5 10:unann6 11:unann7 12:unann8
                   13:unann9 14:unann10 15:unenc 16:unrev

Statistics:      Mean 48.698;  Variance 104.709;  scale 0.465

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution

                              SUMMARIES
                  %
Result           Query
  No.   Score   Match Length DB  ID             Description           Pred. No.
-------------------------------------------------------------------------------
     1   1735    44.4    586   3  QRXLE         estrogen receptor -    0.00e+00
     2   1710    43.8    589   3  QRCHE         estrogen receptor -    6.33e-299
     3   1683    43.1    595   3  QRHUE         estrogen receptor -    1.27e-293
     4   1669    42.8    599   3  QRMSE         estrogen receptor -    7.15e-291
     5   1663    42.6    600   3  QRPTE         estrogen receptor -    1.08e-289
     6   1516    38.8    574   7  A37197        estrogen receptor -    7.15e-261
     7   1503    38.5    535   7  S58224        oestrogen receptor -   2.51e-258
     8   1473    37.7    701  16  S64737        80K estrogen recepto   1.86e-252
     9    807    20.7    433  13  B29345        steroid hormone rece   1.39e-123
    10    799    20.5    433  14  S58087        estrogen receptor re   4.62e-122
    11    800    20.5    521  13  A29345        steroid hormone rece   2.99e-122
    12    681    17.4    923   7  A39596        progesterone recepto   1.[illegible]e-99
    13    671    17.2    922  14  I53280        progesterone recepto   7.61e-98
    14    666    17.1    933   3  QRHUP         progesterone recepti   6.63e-97
    15    662    17.0    930   7  A25923        progesterone recepti   3.74e-96
    16    614    15.7    742   3  QRHUTB        glucocorticoid recep   3.65e-87
    17    614    15.7    777   3  QRHUGA        glucocorticoid recep   3.65e-87
    18    608    15.6    776   3  S44047        glucocorticoid recep   4.82e-86
    19    597    15.3    786   7  A35466        progesterone recepto   5.41e-84
    20    594    15.2    771   7  A54273        glucocorticoid recep   1.96e-83
    21    586    15.0    783   3  A25691        glucocorticoid recep   6.04e-82
    22    579    14.8    132   7  S26595        estrogen receptor -    1.21e-80
    23    577    14.8    795   3  QRRTG         glucocorticoid recep   2.84e-80
    24    571    14.6    899  14  A35895        androgen receptor -    3.70e-79
    25    570    14.6    [illegible]  14  B40494        androgen receptor -    5.67e-79
    26    567    14.5    [illegible]  14  A[illegible]        androgen receptor -    2.04e-78
    27    564    14.4    910  14  A24721        androgen receptor A    7.21e-78
    28    562    14.4    981  14  A41401        mineralocorticoid re   1.73e-77
    29    560    14.3    911  13  B24721        androgen receptor B    4.05e-77
    30    543    13.9    344  13  I51330        androgen receptor -    5.66e-74
    31    544    13.9    758  16  S[illegible]        glucocorticoid recep   3.70e-74
    32    533    13.7    533  13  S37781        retinoid X receptor    3.97e-72
    33    529    13.6    410  14  S26669        retinoic acid recept   2.17e-71
    34    529    13.6    446  14  A34418        H-2 region II bindin   2.17e-71
    35    529    13.6    448  14  D41727        retinoid X receptor    2.17e-71
    36    529    13.6    451  14  A41651        retinoic acid recept   2.17e-71
    37    523    13.4    984  13  A29517        mineralocorticoid re   2.75e-70
    38    514    13.2    466  14  C41727        retinoid X receptor    1.25e-68
    39    514    13.2    467  14  S26668        retinoic acid recept   1.25e-68
    40    514    13.2    467  14  A47278        retinoid X receptor    1.25e-68
    41    513    13.1    462  13  S09592        retinoic acid recept   1.91e-68
    42    502    12.9    563   7  JQ0622        nerve growth factor-   1.99e-66
    43    500    12.8    395  14  I53158        orphan nuclear recep   4.63e-66
    44    499    12.8    488  13  C41977        retinoid receptor al   7.06e-66
    45    501    12.8    538   7  A37201        probable nuclear hor   [illegible]

                              ALIGNMENTS

RESULT     1
ENTRY              QRXLE      #type complete
TITLE              estrogen receptor - African clawed frog
ORGANISM           #formal_name Xenopus laevis #common_name African clawed frog
DATE               28-Feb-1992 #sequence_revision 14-Jul-1994 #text_change
                     15-Jun-1996
ACCESSIONS         A40907
REFERENCE          A40907
   #authors        Weiler, I.J., Lew, D., Shapiro, D.J.
   #journal        Mol. Endocrinol. (1987) 1:355-362
   #title          The Xenopus laevis estrogen receptor: sequence homology with
                     human and avian receptors and identification of multiple
                     estrogen receptor messenger ribonucleic acids.
   #cross-references MUID:90331927
   #accession      A40907
      ##molecule_type mRNA
      ##residues       1-586 ##label WEI
      ##cross-references GB:L20735
COMMENT    The steroid hormones and their receptors are involved in the
             regulation of eukaryotic gene expression and affect cellular
             proliferation and differentiation in target tissues.
COMMENT    In the absence of ligand, steroid hormone receptors are thought to
             be weakly associated with nuclear components; hormone binding
             greatly increases receptor affinity. The hormone-receptor complex
             appears to recognize discrete DNA sequences upstream of
             transcriptional start sites.
CLASSIFICATION   #superfamily estrogen receptor, erbA transforming protein
                     homology
KEYWORDS           DNA binding; nucleus; phosphoprotein; steroid binding;
                     steroid hormone receptor; transcription regulation; zinc
                     finger
FEATURE
   1-174                 #domain amino-terminal #label NHT\
   175-266               #domain DNA binding #status predicted #label DNA\
   179-448               #domain erbA transforming protein homology #label ERBA\
   178-201               #region zinc finger CCCC motif\
   214-237               #region zinc finger CCCC motif\
   251-266               #region nuclear location signal\
   292-544               #domain steroid binding #status predicted #label STE\
   [illegible]           #binding_site zinc (Cys) #status predicted\
   [illegible]           #binding_site zinc (Cys) #status predicted\
   231                   #binding_site phosphate (Ser) (covalent) #status
                           predicted
SUMMARY            #length 586  #molecular-weight [illegible]  #checksum 9991
```

```
  Query Match             44.4%;  Score 1735;  DB 3;  Length 586;
  Best Local Similarity   50.8%;  Pred. No. 0.00e+00;
  Matches  247;  Conservative  112;  Mismatches  107;  Indels   20;  Gaps   16;

Db       70 yssaslsyaassettq-ssaltqlht lnnvppspvvflaklpqlspflhhhqqvpyyl 127
            |:: :: |:::    :: :|::| |            || |: :    :|||:: |: |     |
Qy       43 YPAMIF-YSPAVMNYSIPSNVTNLEGGPGRQTTSPNVLWPTPGHLSPLVVHR-QLSHLYA 100

Db      128 enegqtf avreaapptfyrassdnrrqsqrermssandkappsmestketryeavesdy 186
            |:: :      |            : :|: || :|  ::    || :   | .....||||||
Qy      101 EPQKSPWCEARSLEHTLPVNRETLKRKVSG-NR-CASPVTGP-G--SKRDAHFCAVCSDY 155

Db      187 asqyhygvwsceqckaffkrsiqqhndymcpatnqctidknrrksoqacrlrkcyevgmm 246
            |||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||:
Qy      156 ASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRLRKCYEVGMV 215

Db      247 kqqlrkdrraqllkhkrqkeepeqkmdvdpsslrtasiwvnpsvksmklspvlsltaeq 306
            | | |::| | ||::  ||:  :|| :          |:::    | |: . | . ||  .||
Qy      216 KTCEKEKECYKLVR-FLRSACEQLH-CAGKAKRSGG-HAPRVRELLLDA-LS-PEQ 267

Db      307 liaalmeaeapivysehdstkplsearsmmtlltnladrelvhminwakrvpgfvdltld 366
            |: :|:|||:| |     : |: |::||||  ||:|||:||||||:|||::||||:|:| |
Qy      268 LVLTLLEAEPPHVLISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 326

Db      367 qvhllecawleilmvqliwrsvehpgklsfapnllldrnqgrcveglveifdmlvttatr 426
            ||:|||  |:|:||:||:|||::||||| |||:|:|||::|:||||::||||||:·|··|
Qy      327 QVKLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSR 386

Db      427 trmmrlrqeeficlksiillnsqvytflsstlesledtdlihiildkiidtlvhfmaksg 486
            || ::|:  |::|:|::|||||::| ::::| ::       | |: |: : |:||   :||||
Qy      387 FRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHL-LNAVTDALVWVIAKSG 445

Db      487 lslqqqrrlaqllllilshirhmsnkgmehlysmkcknvvplydlllemldahrihtpkd 546
            :| |||  ||| ||::|||:|| |||||||| :||||||||:||||||||:|| ::   |
Qy      446 ISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNAHVLRGCKS 505

Db      547 ktttqe 552
              |   |
Qy      506 SITGSE 511


RESULT      2
ENTRY                QRCHE      #type complete
TITLE                estrogen receptor - chicken
ORGANISM             #formal_name Gallus gallus #common_name chicken
DATE                 06-Mar-1992 #sequence_revision 14-Jul-1994 #text_change
                       15-Jun-1996
ACCESSIONS           A40914; S07192
REFERENCE            A40914
   #authors          Maxwell, B.L.; McDonnell, D.P.; Conneely, O.M.; Schulz, T.Z.;
                       Greene, G.L.; O'Malley, B.W.
   #journal          Mol. Endocrinol. (1987) 1:25-35
   #title            Structural organization and regulation of the chicken
                       estrogen receptor.
   #cross-references MUID:88318621
   #accession        A40914
      ##status             not compared with conceptual translation
      ##molecule_type mRNA
      ##residues           1-589 ##label MAX
REFERENCE            S07192
   #authors          Krust, A.; Green, S.; Argos, P.; Kumar, V.; Walter, P.;
                       Bornert, J.M.; Chambon, P.
   #journal          EMBO J. (1986) 5:891-897
   #title            The chicken oestrogen receptor sequence: homology with v-erbA
                       and the human oestrogen and glucocorticoid receptors.
   #cross-references MUID:86247578
   #accession        S07192
      ##status             preliminary
      ##molecule_type mRNA
      ##residues           1-256,'P',258-589 ##label KRU
      ##cross-references EMBL:X03805
COMMENT    The steroid hormones and their receptors are involved in the
             regulation of eukaryotic gene expression and affect cellular
             proliferation and differentiation in target tissues.
COMMENT    In the absence of ligand, steroid hormone receptors are thought to
             be weakly associated with nuclear components; hormone binding
             greatly increases receptor affinity. The hormone-receptor complex
             appears to recognize discrete DNA sequences upstream of
             transcriptional start sites.
CLASSIFICATION   #superfamily estrogen receptor, erbA transforming protein
                   homology
KEYWORDS         DNA binding, nucleus, phosphoprotein, steroid binding;
                   steroid hormone receptor, transcription regulation, zinc
                   finger
FEATURE
   1-173                 #domain amino-terminal #label NH2\
   174-265               #domain DNA binding #status predicted #label DNA\
   177-450               #domain erbA transforming protein homology #label ERBA\
   177-200               #region zinc finger CCCC motif\
   215-235               #region zinc finger CCCC motif\
   250-265               #region nuclear location signal\
   294-546               #domain steroid binding #status predicted #label STR\
   179,182,196,199       #binding_site zinc (Cys) #status predicted\
   215,221,231,234       #binding_site zinc (Cys) #status predicted\
   230,299               #binding_site phosphate (Ser) (covalent) #status
                           predicted
SUMMARY          #length 589  #molecular-weight 66773  #checksum 793

  Query Match             43.8%;  Score 1710;  DB 3;  Length 589;
  Best Local Similarity   51.6%;  Pred. No. 6.33e-299;
  Matches  247;  Conservative  112;  Mismatches   99;  Indels   21;  Gaps   14;

Db       69 ygsttlsyaptsesfq-ssslagfhs-lnnvppspvvflqtapqlspfihhhsqqvpyyl 126
             | . |. |.|.  ..: :|::: : :         || |:  |: :|||:: |: |    |
Qy       43 YPAMIF-YSPAVMNYSIPSNVTNLEGGPGRQTTSPNVLWPTPGHLSPLVVHR-QLSHLYA 100

Db      127 eneqqsfqmreaappafyrpssdnrrhsirermsstnekqslsmestketryeavendya 186
            | : ::      :  ::    : :  :: :       ::    |: :  | :::::||||:|||
Qy      101 EPQKSPWCEARSLEHTL-PVNRETLKRKVSGNRCASPVTGP-G--SKRDAHFCAVCSDYA 156

Db      187 sqyhygvwsceqckaffkrsiqghndymcpatnqctidknrrksoqacrlrkcyevgmmk 246
            |||||||||||||||||||||||:|||||||||||||||||||||||||||||||:|
Qy      157 SGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRLRKCYEVGMVK 216

Db      247 gqirkdrrqqrmmkqkrqreeqdsrngeasstelrapilwtspivvkhnkknspalslta 306
             | |::| | |:::: |  :||      | |  :  :|| ::   |::      :  ||| :
Qy      217 CGSPRFRGGYRLVRRPRSADEQLHCAGKAKRSGGHAPRVR-E-LLL-----D-ALS-P 265

Db      307 eqmvsalleaeppivysoydpnrpfneasmmtlltnladrelvhminwakrvpgfvdltl 366
            ||:| :|||||||| |     |: || |||||  ||:|||:||||||:|||::||||:|:|
Qy      266 EQLVLTLLEAEPPHVLIS-RPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSL 324

Db      367 hdqvhllecawleilmiglvwrsmehpgkllfapnlildrnqgkcvegmveifdmllata 426
             |||:|||  |:|:||:||:|||::||||| |||:|:|||::|:||||::||||||||:
Qy      325 FDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLAII 384

Db      427 arfrmmnlqgeefvclksiillnsgvytflsstlksleekdyihrvldkitdtlihlmak 486
            :||| ::||  |::|:|::|||||::| ::::|  :    |    | :|: :||:|: ::||
Qy      385 SRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAH-LLNAVTDALVWVIAK 443

Db      487 sqlslqqqhrrlaqllllilshirhmsnkgmehlynmkcknvvplydlllemldahrlha 545
            ||:| |||  ||| ||::|||:|| |||||||| |||||||||:||||||||:||  |::
Qy      444 SGISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNAHVLRG 502


RESULT      3
ENTRY                QRHUE      #type complete
TITLE                estrogen receptor - human
ALTERNATE_NAMES      hER
ORGANISM             #formal_name Homo sapiens #common_name man
DATE                 28-May-1986 #sequence_revision 28-May-1986 #text_change
                       15-Nov-1996
ACCESSIONS           A94284, A93376, A43021, S27143, S34000, A41925, B41925;
                       A03244; C41925; D41925
REFERENCE            A94284
   #authors          Greene, G.L.; Gilna, P.; Waterfield, M.; Baker, A.; Hort, Y.;
```

```
                  Shine, J.
   #journal       Science (1986) 231:1150-1154
   #title         Sequence and expression of human estrogen receptor
                    complementary DNA.
   #cross-references MUID:86122927
   #accession     A94284
      ##molecule_type mRNA
      ##residues       1-595 ##label GR1
      ##cross-references GB:M12674
REFERENCE         A93376
   #authors       Green, S.; Walter, P.; Kumar, V.; Krust, A.; Bornert, J.M.;
                    Argos, P.; Chambon, P.
   #journal       Nature (1986) 320:134-139
   #title         Human oestrogen receptor cDNA: sequence, expression and
                    homology to v-erb-A.
   #cross-references MUID:86146892
   #accession     A93376
      ##molecule_type mRNA
      ##residues       1-595 ##label GR2
      ##cross-references GB:X03635
REFERENCE         A43021
   #authors       Ponglikitmongkol, M.; Green, S.; Chambon, P.
   #journal       EMBO J. (1988) 7:3385-3388
   #title         Genomic organization of the human oestrogen receptor gene
   #accession     A43021
      ##molecule_type DNA
      ##residues       143-161,205-225,244-264,356-374,'G',376,412-422,447-461,
                        'P',462-467,508-528 ##label PON
REFERENCE         S27140
   #authors       Keaveney, M.; Klug, J.; Gannon, F.
   #journal       DNA Seq. (1992) 2:347-358
   #title         Sequence analysis of the 5' flanking region of the human
                    estrogen receptor gene.
   #accession     S27142
      ##status         preliminary; translation not shown
      ##molecule_type DNA
      ##residues       1-115 ##label KEA
      ##cross-references EMBL:X62462
REFERENCE         S34000
   #authors       Pfeffer, U.; Fecarotta, E.; Castagnetta, L.; Vidali, G.
   #journal       Cancer Res. (1993) 53:741-743
   #title         Estrogen receptor variant messenger RNA lacking exon 4 in
                    estrogen-responsive human breast cancer cell lines.
   #accession     S34000
      ##status         preliminary
      ##molecule_type mRNA
      ##residues       216-254,267-399,'G',401-434 ##label PFE
      ##cross-references EMBL:X73067
      ##note           the authors translated the codon GGG for residue 400 as
                         Val
REFERENCE         A41925
   #authors       Dotzlaw, H.; Alkhalaf, M.; Murphy, L.C.
   #journal       Mol. Endocrinol. (1992) 6:773-785
   #title         Characterization of estrogen receptor variant mRNAs from
                    human breast cancers.
   #cross-references MUID:92293154
   #accession     A41925
      ##molecule_type mRNA
      ##residues       1-214,'ELPTLC' ##label DOT
      ##cross-references NCBIN:106580
      ##experimental_source clone 4; breast cancer
      ##note           sequence has been revised after extraction from NCBI
                         backbone
      ##note           the complete sequence of neither the nucleotide nor the
                         protein is shown in this paper
   #accession     B41925
      ##status         significant sequence differences
      ##molecule_type mRNA
      ##cross-references NCBIN:106597; GB:M69296
      ##experimental_source clone 24; breast cancer
COMMENT    The steroid hormones and their receptors are involved in the
             regulation of eukaryotic gene expression and affect cellular
             proliferation and differentiation in target tissues
```

```
COMMENT    In the absence of ligand, steroid hormone receptors are thought to
             be weakly associated with nuclear components; hormone binding
             greatly increases receptor affinity. The hormone-receptor complex
             appears to recognize discrete DNA sequences upstream of
             transcriptional start sites.
GENETICS
   #gene          GDB:ESR
      ##cross-references GDB:119120
   #map_position 6q25.1-6q25.1
   #introns       151/2, 215/1, 254/1, 366/1, 412/2, 457/1, 516/2
CLASSIFICATION   #superfamily estrogen receptor; erbA transforming protein
                   homology
KEYWORDS         DNA binding; nucleus; phosphoprotein; steroid hormone
                   receptor; transcription regulation; zinc finger
FEATURE
   1-120               #domain amino-terminal #label NH2\
   121-299             #domain DNA binding #status predicted #label DNA\
   183-456             #domain erbA transforming protein homology #label ERBA\
   185-205             #region zinc finger C4C4 motif\
   221-245             #region zinc finger C4C4 motif\
   300-595             #domain steroid binding #status predicted #label STR\
   236,305             #binding_site phosphate (Ser) (covalent) #status
                         predicted
SUMMARY          #length 595  #molecular-weight 66258  #checksum 4317

  Query Match             43.1%;  Score 1683;  DB 3;  Length 595;
  Best Local Similarity   56.6%;  Pred. No. 1.27e-293;
  Matches  226;  Conservative   98;  Mismatches   60;  Indels   17;  Gaps   13;

Db    179 saketrycavcndyasgyhygvwscegckaffkrsiqghndymcpatnqctidknrrksc 237
           | .:.:.||||.|||||||||||||||||||||||||||||||:||||||||||||.|||||:
Qy    142 SKPEAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSC 201

Db    238 qacrlrkcyevgmmkggirkdrrggrmlkhkrqrddgegrgevgsagdmraanlwpsplm 297
           |||||||||||||.| | |..| | |...   |||    | :    :: : :|::     :|
Qy    202 QACRLRKCYEVGMVKCGSRRERCGYRLVR--RQRSADEQ-LHCAGKA-KRSGG--HAP-R 254

Db    298 ikrskknslalsltadqmvsalldaepp-ilyseydptrpfseasmmglltnladrelvh 356
           : |     |  :|::::|: :||:|||| :| |    |: ||:||||'|  '|:|||:||||
Qy    255 V-RELL--LD-ALSPEQLVLTLLEAEPPHVLISR--PSAPFTEASMMMSLTKLADKELVH 308

Db    357 minwakrvpgfvdltlhdqvhllecawleilmiglvwrsmehpgkllfapnllldrnqgk 416
           ||:||||::||||:|:| |||:|||  |:|:'|:||:|||::|| ||:||||:||||::|'
Qy    309 MISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGK 368

Db    417 cvegmveifdmllatssrfrmmnlqgeefvclksiillnsgvytflsstlksleekdhih 476
           ||||::|||||||||.||||  ..||  |:.|.|.:||||||::| ::::|  :  :    :
Qy    369 CVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSRKLA 427

Db    477 rvldkitdtlihlmakagltlqqqhrrlaqllllilshirhmsnkgmehlysmkcknvvpl 536
           ::|: :||:|: ::||:|:: |||  ||| ||::||:|:|| |||||||' :|||'||||:
Qy    428 HLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPV 487

Db    537 ydlllemldahrlha-ptsrggasveetdqshlatagstsshs 578
           |||||||||:|| |:: :|  |:    :::|  :  || :::|
Qy    488 YDLLLEMLNAHVLRGCKSSITGSECSPAEDSK-SKEGSQNPQS 529


RESULT      4
ENTRY            QRMSE       #type complete
TITLE            estrogen receptor - mouse
ORGANISM         #formal_name Mus musculus #common_name house mouse
DATE             06-Mar-1992 #sequence_revision 14-Jul-1994 #text_change
                   15-Jun-1996
ACCESSIONS       A40061
REFERENCE        A40061
   #authors      White, R.; Lees, J.A.; Needham, M.; Ham, J.; Parker, M.
   #journal      Mol. Endocrinol. (1987) 1:735-744
   #title        Structural organization and expression of the mouse estrogen
                   receptor.
   #cross-references MUID:91042558
   #accession    A40061
      ##molecule_type mRNA
      ##residues      1-599 ##label WHI
      ##cross-references GB:M38651
COMMENT    The steroid hormones and their receptors are involved in the
             regulation of eukaryotic gene expression and affect cellular
             proliferation and differentiation in target tissues.
COMMENT    In the absence of ligand, steroid hormone receptors are thought to
             be weakly associated with nuclear components; hormone binding
             greatly increases receptor affinity. The hormone-receptor complex
             appears to recognize discrete DNA sequences upstream of
             transcriptional start sites.
CLASSIFICATION   #superfamily estrogen receptor; erbA transforming protein
                   homology
KEYWORDS         DNA binding; nucleus; phosphoprotein; steroid binding;
                   steroid hormone receptor; transcription regulation; zinc
                   finger
FEATURE
   1-183              #domain amino terminal #label NH2\
   184-275            #domain DNA binding #status predicted #label DNA\
   187-460            #domain erbA transforming protein homology #label ERBA\
   187-210            #region zinc finger CCCC motif\
   223-245            #region zinc finger CCCC motif\
   260-275            #region nuclear location signal\
   304-556            #domain steroid binding #status predicted #label STR\
   189,192,206,209    #binding_site zinc (Cys) #status predicted\
   225,231,241,244    #binding_site zinc (Cys) #status predicted\
   240,309            #binding_site phosphate (Ser) (covalent) #status
                        predicted
SUMMARY          #length 599  #molecular-weight 66954  #checksum 4749

  Query Match             42.8%;  Score 1669;  DB 3;  Length 599;
  Best Local Similarity   51.7%;  Pred. No. 7.15e-291;
  Matches  246;  Conservative  106;  Mismatches  104;  Indels  20;  Gaps  14;

Db     108 spsplmllhpppqlspflhphaqqvpyylenepsay-avrdtqppafyrsnsdnrrqngr 166
            : || :|    | :||| |  | |     | | : |:     |            :: :|: :|
Qy      73 TTSPNVLWPTPGHLSP-LVVHRQLSHLYAEPQKSPWCEARSLEHTLPVNRETLKRKVSG- 130

Db     167 erlsssnekgsmlmesaketryc-avcndyasqyhygvwsceqckaffkrsiqghndymcp 226
            :|   :|     |       | :::::||||:|||||||||||||||||||||||||||||:||
Qy     131 NR-CASPVTGP---GSKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICP 186

Db     227 atnqctidknrrksoqacrlrkcyevqmmkggirkdrrggrmlkhkrqrddleqrnemqa 286
           |||||||||||||||||||||||||||||:| | |::| | |:::  |||    : :    |
Qy     187 ATNQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVR--RQRSA-DEQLH-CA 242

Db     287 sqdmraanlwpsplvikhtkknspalsltadqmvsalldaeppmiyseydpsrpfseasm 346
           :  :|::     :| | :      :  |||  ::|:| :||:|||| :      || ||:||||
Qy     243 GKAKRSGG--HAPRV-RELLLD--ALS--PEQLVLTLLEAEPPHVLIS-RPSAPFTEASM 294

Db     347 mqlltnladrelvhminwakrvpqfqdlnlhdqvhllecawleilmiglvwrsmehpgkl 406
           |   ||:|||:||||||.|||..|||  .|.| |||.|||   |.|.||.||.|||..|||||
Qy     295 MMSLTKLADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKL 354

Db     407 lfapnllldrnqgkcveqmveifdmllatssrfrmmnlqgeefvclksiillnsgvytfl 466
           :|||:|:|||::||||||::||||||||||:||||  ::||   |::|:|:·||||||··|  ··
Qy     355 IFAPDLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLV 414

Db     467 sstlksleekdhihrvldkitdtlihlmakagltlqqqhrrlaqllliilshirhmsnkgm 526
           ::|  :  :      : ::|: :||:|: ::||:|:: |||  ||| ||::|||:|| |||||
Qy     415 TATQDA-DSSRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGM 473

Db     527 ehlysmkcknvvplydlllemldahrlhapasrmqvppeepsqtqlattsstsahs 582
           |||  |||||| ||:||||||||:|| |::   |  :  :   |::    :  :| ::.|
Qy     474 EHLLNMKCKNVVPVYDLLLEMLNAHVLRGCKSSITCSECSPAEDSKSKEGSQNPQS 529


RESULT       5
ENTRY               QRRTE       #type complete
TITLE               estrogen receptor - rat
ORGANISM            #formal_name Rattus norvegicus #common_name Norway rat
DATE                30-Sep-1991 #sequence_revision 14-Jul-1994 #text_change
                      15-Jun-1996
ACCESSIONS          S07379; S16731
```

```
REFERENCE        S07379
   #authors      Koike, S.; Sakai, M.; Muramatsu, M.
   #journal      Nucleic Acids Res. (1987) 15:2499-2513
   #title        Molecular cloning and characterization of rat estrogen
                   receptor cDNA.
   #cross-references MUID:87174780
   #accession    S07379
      ##molecule_type mRNA
      ##residues      1-600 ##label KOI
      ##cross-references EMBL:Y00102
REFERENCE        S16731
   #authors      Maggi, A.M.A.
   #submission   submitted to the EMBL Data Library, June 1991
   #accession    S16731
      ##molecule_type mRNA
      ##residues      1-487,'T',489-600 ##label MAG
      ##cross-references EMBL:X61098
COMMENT    The steroid hormones and their receptors are involved in the
             regulation of eukaryotic gene expression and affect cellular
             proliferation and differentiation in target tissues.
COMMENT    In the absence of ligand, steroid hormone receptors are thought to
             be weakly associated with nuclear components; hormone binding
             greatly increases receptor affinity. The hormone-receptor complex
             appears to recognize discrete DNA sequences upstream of
             transcriptional start sites.
CLASSIFICATION   #superfamily estrogen receptor; erbA transforming protein
                   homology
KEYWORDS         DNA binding; nucleus; phosphoprotein; steroid binding;
                   steroid hormone receptor; transcription regulation; zinc
                   finger
FEATURE
   1-184              #domain amino-terminal #label NH2\
   185-276            #domain DNA binding #status predicted #label DNA\
   188-461            #domain erbA transforming protein homology #label ERBA\
   188-211            #region zinc finger CCCC motif\
   224-246            #region zinc finger CCCC motif\
   261-276            #region nuclear location signal\
   305-557            #domain steroid binding #status predicted #label STR\
   190,193,207,210    #binding_site zinc (Cys) #status predicted\
   226,232,242,245    #binding_site zinc (Cys) #status predicted\
   241,310            #binding_site phosphate (Ser) (covalent) #status
                        predicted
SUMMARY          #length 600  #molecular-weight 67029  #checksum 4599

  Query Match             42.6%;  Score 1663;  DB 3;  Length 600;
  Best Local Similarity   51.1%;  Pred. No. 1.08e-289;
  Matches  243;  Conservative  110;  Mismatches  103;  Indels  20;  Gaps  14;

Db     109 spsplmllhppphvspflhphahqvpyylenepsay-avrdtqppafyrsnsdnrrqnqr 167
            : || :|    | |:|| |  | :     | | : |:     |            :: :|: :|
Qy      73 TISPNVLWPTPGHLSP-LVVHRQLSHLYAEPQKSPWCEARSLEHTLPVNRETLKRKVSG- 130

Db     168 erlsssseknqmimesaketrycavcndyasqyhygvwsceqckaffkrsiqghndymcp 227
            :|   :|:   |       | :::::||||:|||||||||||||||||||||||||||||:||
Qy     131 NR-CASPVTGP---GSKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICP 186

Db     228 atnqctidknrrksoqacrlrkcyevqmmkggirkdrrggrmlkhkrqrddleqrnemqt 287
           |||||||||||||||||||||||||||||:| | |::| | |:::: |  |: : :    |
Qy     187 ATNQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQRSADE-QLHCA-GK 244

Db     288 sqdmraanlwpsplvikhtkknspalsltadqmvsalldaepplIyseydpsrpfseasm 347
           :  :|::     :| | :      :  |||  ::|:| :||:|||| :      || ||:||||
Qy     245 A--KRSGG--HAPRV-RELLLD--ALS--PEQLVLTLLEAEPPHVLIS-RPSAPFTEASM 294

Db     348 mglltnladrelvhminwakrvpqfqdlnlhdqvhllecawleilmiglvwrsmehpgkl 407
           |   ||:|||:||||||:|||::|||  :|:| |||:|||   |:|:||:||:|||::|||||
Qy     295 MMSLTKLADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKL 354

Db     408 lfapnllldrnqgkcveqmveifdmllatssrfrmmnlqgeefvclksiillnsgvytfl 467
            ||| | |||   ||||||..|||||||||.||||  ..||   |..|.|..|||||::| ::
Qy     355 IFAPDLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLV 414

Db     468 sstlksleekdhihrvldkindtlihlmakagltlqqqhrrlaqllliilshirhmsnkgm 527
```

```
              ::|   : :     : ::|: . |.|. ..||.|.: |||  ||| ||::||:|| |||||
Qy    415 TATQDA-DSSRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGM 473

Db    528 ehlynmkckavvplydlllemldahrlhapasrmgvppeepsqsqltttsstsahs 583
          |:| |||||||||.|||||||||.|| |:.   | .   .    |..    .  .| ...|
Qy    474 EHLLNMKCKNVVPVYDLLLEMLNAHVLRGCKSSITGSECSPAEDSKSKEGSQNPQS 529


RESULT     6
ENTRY              A37197     #type complete
TITLE              estrogen receptor - rainbow trout
ORGANISM           #formal_name Oncorhynchus mykiss #common_name rainbow trout
DATE               31-Jan-1992 #sequence_revision 31-Jan-1992 #text_change
                     03-Jun-1996
ACCESSIONS         A37197; A40070
REFERENCE          A37197
   #authors        Pakdel, F.; Le Gac, F.; Le Goff, P.; Valotaire, Y.
   #journal        Mol. Cell. Endocrinol. (1990) 71:195-204
   #title          Full-length sequence and in vitro expression of rainbow trout
                     estrogen receptor cDNA.
   #cross-references MUID:91006824
   #accession      A37197
      ##status          preliminary
      ##molecule_type mRNA
      ##residues        1-574 ##label PAK
      ##cross-references GB:M31559
REFERENCE          A40070
   #authors        Pakdel, F.; Le Guellec, C.; Vaillant, C.; Le Roux, M.G.;
                     Valotaire, Y.
   #journal        Mol. Endocrinol. (1989) 3:44-51
   #title          Identification and estrogen induction of two estrogen
                     receptors (ER) messenger ribonucleic acids in the rainbow
                     trout liver: sequence homology with other ERs.
   #cross-references MUID:89127284
   #accession      A40070
      ##status          preliminary
      ##molecule_type mRNA
      ##residues        150-574 ##label PA2
      ##cross-references GB:M31559
CLASSIFICATION     #superfamily estrogen receptor; erbA transforming protein
                     homology
KEYWORDS           DNA binding; nucleus; steroid hormone receptor; zinc finger
FEATURE
   145-421              #domain erbA transforming protein homology #label ERBA\
   147-167              #region zinc finger\
   183-207              #region zinc finger
SUMMARY            #length 574  #molecular-weight 62871  #checksum 1057

  Query Match             38.8%;  Score 1516;  DB 7;  Length 574;
  Best Local Similarity   56.3%;  Pred. No. 7.15e-261;
  Matches  225;  Conservative   78;  Mismatches  75;  Indels  22;  Gaps  17;

Db    143 etrycavcsdfasgyhygvwscegckaffkrsiqghndymcpatnqctmdrnrrkscqac 202
          ::::|||||:|||||||||||||||||||||||||||||||:||||||| | |||||||||
Qy    145 DAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQAC 204

Db    203 rlrkcyevgmvkgglrkdrgg-rvlrkdkryrgpagdrekptvtwstgqrpgdggrnsss 261
          ||||||||||||| | |::| | |..|:. |   .| | |     : . . |.  ||.  .
Qy    205 RLRKCYEVGMVKCGSRRERCGYRLVRRQ-R---SA-D-EQLHCA-GKAKRS--GGH--AP 253

Db    262 slggggwrgpritmppeqvlfllqggtp-alcsrqkvarpytevtmmtlltsmadkelv 320
           :          : .. | . || ||::. |  | ||   | |.|| .||  || :||||||
Qy    254 RVREL--LLDA-IS-PEQLVLTLLEAEPPHVLISRPS-A-PFTEASMMMSLTKLADKELV 307

Db    321 hmiawakkvpgfqelslhdqvqllesswlevlmigliwrsihcpgklifaqdlildrseg 380
          |||:|||||:|'| |||| |||:|||| |:|||||:||:|||||  ||||||| ||:||| ||
Qy    308 HMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEG 367

Db    381 dcvegmaeifdmllatvsrfrmlklkpeefvclkailllnsgafsfcsnsveslhnssav 440
           ||||: ||||||||| |||  |||   |::|:||:||||:: ... . . .. .|  .
Qy    368 KCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSRKL 427

Db    441 esmldnitdallhhishsgasvqqqprrqaqlllllshirhmsnkgmehlysikcknkvp 500
```

```
              :|: :||||:  |: || | |||: | | ||:||||:|| ||||||| ::||| ||
Qy    427 AHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVP 486

Db    501 lydlllemldghrlqapgkvaqageqtegpstttttstgs 540
          :||||||||::| |::  | : :| :  :': : :  ||
Qy    487 VYDLLLEMLNAHVLRGC-KSSITGSEC-SPAEDSKSKEGS 524


RESULT     7
ENTRY              S58224     #type fragment
TITLE              oestrogen receptor - Atlantic salmon (fragment)
ORGANISM           #formal_name Salmo salar #common_name Atlantic salmon
DATE               13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change
                     26-Apr-1996
ACCESSIONS         S58224
REFERENCE          S58224
   #authors        Rogers, S.A.; Llewellyn, L.; Ramsurn, V.P.; Sweeney, G.E.;
                     Wigham, T.
   #submission     submitted to the EMBL Data Library, July 1995
   #accession      S58224
      ##status          preliminary
      ##molecule_type mRNA
      ##residues        1-535 ##label ROG
      ##cross-references EMBL:X89959
CLASSIFICATION     #superfamily erbA transforming protein homology
FEATURE
   102-382              #domain erbA transforming protein homology #label ERBA
SUMMARY            #length 535  #checksum 9963

  Query Match             38.5%;  Score 1503;  DB 7;  Length 535;
  Best Local Similarity   55.7%;  Pred. No. 2.51e-258;
  Matches  224;  Conservative   75;  Mismatches  80;  Indels  23;  Gaps  16;

Db    101 etrycavcsdfasgyhygvwscegckaffkrsiqghndymcpatnqctidrnrrkscqac 160
          ::::||||||:|||||||| ||||||||||||||||||||||:|,||||||:':'|||||||'
Qy    145 DAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQAC 204

Db    161 rlrkcyevgmvkgglrkdrgg-rvlrkdkrycgpagdrekpygdlehrtappqdggrnss 219
          |||||||||||| | |::| | |::|:: |    :| : :    :     : |  : ||: :
Qy    205 RLRKCYEVGMVKCGSRRERCGYRLVRRQ-R---SADE-QLHCAG---K-AK-RSGGHAPR 254

Db    220 ssslsggggwrgpritmppeqvlfllqgaeppalcsrqkvarpytevtmmtlltsmadke 279
             |          : :: | : || ||::  |  | ||   | |:|| :||  || :||||:
Qy    255 VRELL----L-DA-IS-PEQLVLTLLEAEPPHVLISRPS-A-PFTEASMMMSLTKLADKE 305

Db    280 lvhmiawakkvpgfqelslhdqvqllesswlevlmigliwrsihcpgklifaqdlildrs 339
          |||||.||||:||| |||| |||.|||| |.||||.||:||||  ||||||| ||:|:|
Qy    306 LVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRD 365

Db    340 egdcvegmaeifdmllatvsrfrmlklkpeefvclkailllnsgafsfcsnsveslhnss 399
          || ||||: ||||||||| |||| |||   |::|:||:||||||: ::: : : ::  :'
Qy    366 EGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSR 424

Db    400 avesmldnitdallhhishsgasvqqqprrqaqlllllshirhmsnkgmehlysikckrk 459
            :  :|: :||||:  |: || | |||: |   ||:||||:|| ||||||'|| ::||'|
Qy    425 KLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNV 484

Db    460 vplydlllemldghrlqspgkvaqageqtegpstttttstgs 501
          ||:|||||||||::| |::  | : :| :  :|: : :  ||
Qy    485 VPVYDLLLEMLNAHVLRGC-KSSITGSEC-SPAEDSKSKEGS 524


RESULT     8
ENTRY              S64737     #type complete
TITLE              80K estrogen receptor - human
ORGANISM           #formal_name Homo sapiens #common_name man
DATE               28-Oct-1996 #sequence_revision 28-Oct-1996 #text_change
                     28-Oct-1996
ACCESSIONS         S64737
REFERENCE          S64737
   #authors        Pink, J.J.; Wu, S.Q.; Wolf, D.M.; Bilimoria, M.M.; Jordan,
                     V.C.
   #journal        Nucleic Acids Res. (1996) 24:962-969
   #title         A novel 80 kDa human estrogen receptor containing a
                    duplication of exons 6 and 7.
   #accession     S64737
      ##status          preliminary
      ##residues        1-701 ##label PIN
      ##cross-references EMBL:U47678
SUMMARY            #length 701  #molecular-weight 73430  #checksum 4951

  Query Match             37.7%;  Score 1473;  DB 16;  Length 701;
  Best Local Similarity  56.9%;  Pred. No. 1.86e-252;
  Matches  197;  Conservative   35;  Mismatches  49;  Indels  15;  Gaps  11;

Db     178 saketrycavcndyasgyhygvwscegckaffkrsiqghndymcpatnqctidknrrksc 237
           | :::::||||:||||||||||||||||||||||||||||||||:||||||||||||||||
Qy     142 SKPDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSC 201

Db     238 qacrlrkcyevqmmkqgirkdrrggrmlkhkrqrddgegrgevgsagdmraanlwpsplm 297
           ||||||||||||||:| | |::| | |:::  |||    | :    :: : :|::     :|
Qy     202 QACRLRKCYEVGMVKCGSRRERCGYRLVR--RQRSADE-QLHCAGKA KRSGG--HAP-R 254

Db     298 ikrskknslalsltadqmvsalldaepp ilyseydptrpfseasmmglltnladrelvh 356
           : |      |  :|:::|:| :||:|||| :| |    |: ||:||||| ||:|||:||||
Qy     255 V-RELL--LD-ALSPEQLVLTLLEAEPPHVLISR--PSAPFTEASMMMSLTKLADKELVH 308

Db     357 minwakrvpqfvdltlhdqvhllecawleilmiglvwrsmehpvkllfapnllldrnqgk 416
           ||:|||::||||:|:| |||:|||   |.|.||:||.|||..|| ||.|||.|:|||..||
Qy     309 MISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGK 368

Db     417 cvegmveifdmllatssrfrmmnlqgeefvclksiillnsgvytflsstlksleekdhih 476
           ||||::|||||||||:|||| ::||  |::|:|::|||||::| ::..|  .  .    .
Qy     369 CVEGILEIFDMLLATISRFRELKLQHKEYLCVKAMILLNSSMYPLVIAIQDA-DSSRKLA 427

Db     477 rvldkitdtlihlmakaqltlqqqhqrlaqllilshirhmrnqgk 522
           ::|: :||:|: ::||:|:: |||  ||| ||::|||:||  | |:
Qy     428 HLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGM 473


RESULT     9
ENTRY              B29345      #type complete
TITLE              steroid hormone receptor ERR2 precursor - human
ORGANISM           #formal_name Homo sapiens #common_name man
DATE               15-Dec-1988 #sequence_revision 15-Dec-1988 #text_change
                     03-Jun-1996
ACCESSIONS         B29345
REFERENCE          A29345
   #authors        Giguere, V.; Yang, N.; Segui, P.; Evans, R.M.
   #journal        Nature (1988) 331:91-94
   #title          Identification of a new class of steroid hormone receptors.
   #cross-references MUID:88122546
   #accession      B29345
      ##molecule_type mRNA
      ##residues        1-433 ##label GIG
      ##cross-references GB:X51417; GB:Y00290
CLASSIFICATION   #superfamily erbA transforming protein homology
KEYWORDS           DNA binding; zinc finger
FEATURE
   101-352                #domain erbA transforming protein homology #label ERRA\
   103-123                #region zinc finger\
   139-163                #region zinc finger
SUMMARY            #length 433  #molecular-weight 48287  #checksum 1133

  Query Match             20.7%;  Score 807;  DB 13;  Length 433;
  Best Local Similarity  39.5%;  Pred. No. 1.39e-123;
  Matches  140;  Conservative   84;  Mismatches 108;  Indels  22;  Gaps  14;

Db     101 rlclvcgdiasgyhygvascеackaffkrtiqgnieyscpatneceitkrrrkscqacrf 160
           ::| ||:| ||||||||| |||:||||||||:|||: :| ||||||:| | | ||||||||:
Qy     147 HFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRL 206

Db     161 mkclkvgmlkeqvrldrv--r-qgrqky-krrldsenspylslqisppakkpltkivs-- 214
            ||  |||:| | | :|    |   ||:     :|        :  |   :|  :  |    :|
Qy     207 RKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPE 266

Db     215 yl l-vaeqpdklyam-ppddvpeqdlikalttlsdladrelvfllswakhlpgfasltlad 271
           | | : |::      : :      |   :       : :|  |||:||| :||||| |||| :|:| |
Qy     267 QLVLILLEAEPPHVLISRPSAPFTEASMMMSLIKLADKELVHMISWAKKIPGFVELSLFD 326

Db     272 qmsllqsawmeilllgivyrslpyddklayaedyimdeehsrlv-glleiyrailqlvrr 330
           |: ||:| |||:|::|:: ||:     || :| | ::| : :: | |:||:  :!     |
Qy     327 QVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATISR 386

Db     331 ykklkvekeefvmlkalalansdsmy--i---enleavqklqdllh---ealq-dyels- 380
           :: ||::  |:: :||: | || |||   :     :: :: :||  ||:     :||       |
Qy     387 FRELKLQHKEYLCVKAMILLNS-SMYPLVIAIQDADSSRKLAHLLNAVIDALVWVIAKSG 445

Db     381  qhheeprragkllltlpllrqtaakavqhfysvklqgkvpmhklflemleakv 433
              :::: | ::||: |: :|::: |:::|: ::|     ||:  |:||||:| |
Qy     446 ISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNAHV 499


RESULT    10
ENTRY              S58087      #type complete
TITLE              estrogen receptor related protein - mouse
ORGANISM           #formal_name Mus musculus #common_name house mouse
DATE               13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change
                     26-Apr-1996
ACCESSIONS         S58087
REFERENCE          S58087
   #authors        Pettersson, K.; Svensson, K.; Mattsson, R.; Carlsson, B.;
                     Ohlsson, R.; Berkenstam, A.
   #submission     submitted to the EMBL Data Library, July 1995
   #description    Expression of a novel member of estrogen response
                     element-binding nuclear receptors is restricted to the
                     early stages of chorion formation during mouse
                     embryogenesis.
   #accession      S58087
      ##status          preliminary
      ##molecule_type mRNA
      ##residues        1-433 ##label PET
      ##cross-references EMBL:X89594
CLASSIFICATION   #superfamily erbA transforming protein homology
FEATURE
   101-352                #domain erbA transforming protein homology #label ERRA
SUMMARY            #length 433  #molecular-weight 48400  #checksum 1075

  Query Match             20.5%;  Score 799;  DB 14;  Length 433;
  Best Local Similarity  38.7%;  Pred. No. 4.62e-122;
  Matches  137;  Conservative   85;  Mismatches 110;  Indels  22;  Gaps  15;

Db     101 rlclvcgdiasgyhygvascеackaffkrtiqgnieyncpatneceitkrrrkscqacrf 160
           ::| ||:| ||||||||| |||:||||||||:|||: :| ||||||:| | | ||||||||:
Qy     147 HFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRL 206

Db     161 mkclkvgmlkeqvrldrv--r-qgrqky-krrldsenspylslqisppakkpltkivs-n 215
            ||  |||:| | | :|    |   ||:     :|        :  :   :|  :  |    :| :
Qy     207 RKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPE 266

Db     216 -l-lgv-eqdklyam ppadlpeqdikalttlselsdrelvfllnwakhlpgfpsltlad 271
           | | . | .      . .      |   .     . .| |||:||| :|:||| ||||  |:| |
Qy     267 QLVLILLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 326

Db     272 qmsllqsawmeililgivyrslpyddklayaedyimdeehsrlv-qlldlyrailqlvrr 330
           |: ||:| |||:|::|:: ||:     || :| | ::| : :: | |:||:  :|     |
Qy     327 QVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSR 386

Db     331 ykklkvekeefmilkalalansdsmy--i---enleavqklqdllh---ealq-dyels- 380
           :: ||::  |:: :||: | || |||   :     :: :: :||  ||:     :||       |
Qy     387 FRELKLQHKEYLCVKAMILLNS-SMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKSG 445

Db     381  qrheeprragkllwtlpllrqtaakavqhfysvklqgkvpmhklflemleakv 433
              :::: | ::||  |: :|::: |:::|: ::|     ||:  |:||||:| |
Qy     446 ISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNAHV 499


RESULT    11
ENTRY              A29345      #type complete
```

```
TITLE             steroid hormone receptor ERR1 precursor - human
ALTERNATE_NAMES   estrogen-related receptor
ORGANISM          #formal_name Homo sapiens #common_name man
DATE              15-Dec-1988 #sequence_revision 15-Dec-1988 #text_change
                    03-Jun-1996
ACCESSIONS        A29345; A49074
REFERENCE         A29345
   #authors       Giguere, V.; Yang, N.; Segui, P.; Evans, R.M.
   #journal       Nature (1988) 331:91-94
   #title         Identification of a new class of steroid hormone receptors
   #cross-references MUID:88122546
   #accession     A29345
      ##molecule_type mRNA
      ##residues      1-521 ##label GIG
      ##cross-references EMBL:X51416; EMBL:Y00290
REFERENCE         A49074
   #authors       Wiley, S.R.; Kraus, R.J.; Zuo, F.; Murray, E.E.; Loritz, K.;
                    Mertz, J.E.
   #journal       Genes Dev. (1993) 7:2206-2219
   #title         SV40 early-to-late switch involves titration of cellular
                    transcriptional repressors
   #accession     A49074
      ##status        preliminary
      ##molecule_type protein
      ##residues      166-169 'X' 171-173 ##label WIL
CLASSIFICATION    #superfamily erbA transforming protein homology
KEYWORDS          DNA binding; nucleus; zinc finger
FEATURE
   174-404              #domain erbA transforming protein homology #label ERBA\
   176-196              #region zinc finger\
   212-236              #region zinc finger
SUMMARY           #length 521  #molecular-weight 55742  #checksum 5840

  Query Match             20.5%;  Score 800;  DB 13;  Length 521;
  Best Local Similarity   38.9%;  Pred. No. 2.99e-122;
  Matches  141;  Conservative   74;  Mismatches 128;  Indels  10;  Gaps   7;

Db    174 rlclvcgdvasgyhygvasceackaffkrtiqgsieyscpasneceitkrrrkacqacrf 233
           ::| |||:| |||||||| |||.||||||| |||    | ||| | | | | ||| |||||
Qy    147 HFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRL 206

Db    234 tkclrvgmlkegvrldrvrggrqkykrrpevdplpfpg-p-fpagplavaggprktaapv 291
            ||  |||:| | | :| | |    ..|.  . |   .' :  ..|  .        | .
Qy    207 RKCYEVGMVKGGSRRERC-GYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSP 265

Db    292 palvshllvvepeklyampdpagpdghlpavatlcdlfdreivvtiswaksipgfsslsl 351
           : ||  ||  ||  :  :: |::|     : : :|  | |:|:| ||||| |||| |||
Qy    266 EQLVLTLLEAEPPHVL-ISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSL 324

Db    352 sdqmsvlqsvwmevlvlgvaqrslplqdelafaedlvld-eegaraaglgelgaallqlv 410
            ||: :|:| |||||::|:  ||.      | || ||||| .||       |. |.   ||
Qy    325 FDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATT 384

Db    411 rrlqalrlereeyvllkalalansdsvhiedaeaveqlreallh--eal---leyeagra 465
            |:: |:|:: :|: :||: | ||     :     :     | :| |   :|:    |  :::
Qy    385 SRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKS 444

Db    466 gpggaerrragrllltlpllrqtagkvlahfygvklegkvpmhklflemlea 519
           | ::  :  | : ||: |: :|::: | : |:  :|     ||:  |:||||:|
Qy    445 GISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNA 497

RESULT   12
ENTRY             A39596      #type complete
TITLE             progesterone receptor B form - mouse
ORGANISM          #formal_name Mus musculus #common_name house mouse
DATE              20-Mar-1992 #sequence_revision 20-Mar-1992 #text_change
                    03-Jun-1996
ACCESSIONS        A39596
REFERENCE         A39596
   #authors       Schott, D.R.; Shyamala, G.; Schneider, W.; Parry, G.
   #journal       Biochemistry (1991) 30:7014-7020
   #title         Molecular cloning, sequence analyses, and expression of
```

```
                    complementary DNA encoding murine progesterone receptor.
   #cross-references MUID:91299759
   #accession     A39596
      ##status        preliminary
      ##molecule_type mRNA
      ##residues      1-923 ##label SCH
      ##cross-references GB:J05333
CLASSIFICATION    #superfamily progesterone receptor; erbA transforming protein
                    homology
KEYWORDS          DNA binding; nucleus; steroid hormone receptor; zinc finger
FEATURE
   555-819              #domain erbA transforming protein homology #label ERBA\
   557-577              #region zinc finger\
   593-617              #region zinc finger
SUMMARY           #length 923  #molecular-weight 99073  #checksum 6627

  Query Match             17.4%;  Score 681;  DB 7;  Length 923;
  Best Local Similarity   30.8%;  Pred. No. 1.00e-99;
  Matches  123;  Conservative  112;  Mismatches 139;  Indels  27;  Gaps  27;

Db    491 aapgtaaapaiyqplg-lnglpqlg-y-qaavlkdslpqvyppylnylrpdseasqspqy 547
            ::||  ::  : : |: | : || :     :|  |   :|| |     |:  :  ::    ||
Qy     81 PTPGHLSPLVVHRQLSHLYAEPQKSPWCEARSLEHTLP-VNRETLK-RKVSGNRCASPVT 138

Db    548 gfdslpq-kiclicgdeasgchygvltcgsckvffkramegqhnylcagrndcivdkirr 606
            |  |  :  .| .|.| ||| |||| .| .|| ||||...|...|.|.: |:| :|| ||
Qy    139 GPGSKPEAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRR 198

Db    607 kncpacrlrkccqagmvl-ggrkfk-kfnkvrvmraldavalpqsvglpnesqalsqrit 664
            |:| ||||'|| : |||  |:|: :  :  ||  |: |    ' :  | :: | : : |:
Qy    199 KSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQRSADE-QL-HCAGKAKRSGGHAPRVRE 256

Db    665 fspnqeiqlvp-plinllmsiepdviyaghdntkpdtsssllTslnqlgerqllsvvkws 723
                .  | |  |.  |.  ||   .   .. .. | | .|:: ||  |:::|: :: ':
Qy    257 -ELLLDA-LSPEQLVLTLLEAEPPHVLISRPSA-PFTEASMMMSLTKLADKELVHMISWA 313

Db    724 kslpgfrnlhiddqitliqyswmslmvfglgwrsykhvsgqmlyfapdlilne-q-r-mk 780
            | :|||  .| : ||. |..  || ::. || ||| | .| :. |||||:|: :: : :
Qy    314 KKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVE 371

Db    781 e-lsfyslcltmwqipqefvklqvtheeflcmkvllllnt-i-pleglrsqsqfeemrss 837
             | .. . | .        |  |  . | |.|||.| ..|||. : ||      ::: : :
Qy    372 GILEIFDM-L-L-ATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAH 428

Db    838 yirelikaig-lrqkgvvpssqrfyqltklldslhdlvkq 876
                  .     |   |:     |     |  ||  |   | .
Qy    429 LLNAVTDALVWVIAKSGISSQQQSMRLANLL-MLLSHVRH 467

RESULT   13
ENTRY             I53280      #type complete
TITLE             progesterone receptor - rat
ORGANISM          #formal_name Rattus norvegicus #common_name Norway rat
DATE              02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change
                    02-Aug-1996
ACCESSIONS        I53280
REFERENCE         I53280
   #authors       Park-Sarge, O.K.; Mayo, K.E.
   #journal       Endocrinology (1994) 134:709-718
   #title         Regulation of the progesterone receptor gene by gonadotropins
                    and cyclic adenosine 3',5'-monophosphate in rat granulosa
                    cells.
   #cross-references MUID:94130817
   #accession     I53280
      ##status        preliminary; translated from GB/EMBL/DDBJ
      ##molecule_type mRNA
      ##residues      1-923 ##label RES
      ##cross-references GB:L16922; NID:g463282; CDS_PID:g463283
SUMMARY           #length 923  #molecular-weight 99407  #checksum 9843

  Query Match             17.2%;  Score 671;  DB 14;  Length 923;
  Best Local Similarity   32.6%;  Pred. No. 7.61e-98;
  Matches  109;  Conservative   94;  Mismatches 109;  Indels  21;  Gaps  21;
```

```
Db    556 iclicqdeasgchygvltcgsckvffkramegqhnylcagrndcivdkirrkncpacrlr 615
          :| :|:| ||| |||| :| :|| ||||:::|:::|:|:: |:| :|| |||:| |||||
Qy    148 FCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCIIDKNRRKSCQACRLR 207

Db    616 kccqagmvl qgrkfk-kfnkvrvmraldgvalpqsvafpnesqtlqqrltfspnqeiql 673
          || : |||  |:|: :  :  ||  |: |   | :  : :: |    : |:        : |
Qy    208 KCYEVGMVKCGSRRERCGYRLVRRQRSADE-QL-HCAGKAKRSGGHAPRVR-ELLLSA-L 263

Db    674 vp-plinllmsiepdvvyaghdatkpdtssslltslnqlgerqllsvvkwskslpgfrnl 732
           | |:  |:  ||  |  :: :. | | .|.. ||  |::::|: :: |:| :||| :|
Qy    264 SPEQLVLTLLEAEPPHVLISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVEL 322

Db    733 hiddqitlliqyswmslmvfglqwrsykhvsgqmlyfapdliln-eq-r-mke-lsfyslc 788
           : ||: |::  || ::: || |||  | :| .. |||||.|. .. .  .  | .. .
Qy    323 SLFDQVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDM- 379

Db    789 ltmwqipqefvklqvtheeflcmkvllllnt-i-pleglrsqsqfeemrssyirelikai 846
          | :      |  |  : | |:||:| ::|||: : ||      :::   : :  :  :  |:
Qy    380 L-LATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDAL 437

Db    847 q-lrqkqvvpssqrfyqltklldslhdlvkq 876
            :  |: ::| |:  :|::||  |   |::
Qy    438 VWVIAKSGISSQQQSMRLANLL-MLLSHVRH 467


RESULT    14
ENTRY              QRHUP       #type complete
TITLE              progesterone receptor form B - human
ALTERNATE_NAMES    hPR
CONTAINS           progesterone receptor form A
ORGANISM           #formal_name Homo sapiens #common_name man
DATE               30-Jun-1987 #sequence_revision 18-Nov-1994 #text_change
                     15-Nov-1996
ACCESSIONS         S09971; S12464; A03245
REFERENCE          S09971
   #authors        Kastner, P.; Krust, A.; Turcotte, B.; Stropp, U.; Tora, L.;
                     Gronemeyer, H.; Chambon, P
   #journal        EMBO J. (1990) 9:1603-1614
   #title          Two distinct estrogen-regulated promoters generate
                     transcripts encoding the two functionally different human
                     progesterone receptor forms A and B.
   #cross-references MUID:90228361
   #accession      S09971
      ##molecule_type mRNA
      ##residues        1-933 ##label KAS
      ##cross-references EMBL:X51730
REFERENCE          S12464
   #authors        Kastner, P.
   #submission     submitted to the EMBL Data Library, February 1990
   #accession      S12464
      ##molecule_type mRNA
      ##residues        1-343,'T',345-933 ##label KA2
      ##cross-references EMBL:X51730
REFERENCE          A03245
   #authors        Misrahi, M.; Atger, M.; D'Auriol, L.; Loosfelt, H.; Meriel,
                     C.; Fridlansky, F.; Guiochon-Mantel, A.; Galibert, F.;
                     Milgrom, E.
   #journal        Biochem. Biophys. Res. Commun. (1987) 143:740-748
   #title          Complete amino acid sequence of the human progesterone
                     receptor deduced from cloned cDNA.
   #cross-references MUID:87184565
   #accession      A03245
      ##molecule_type mRNA
      ##residues        1-225,'G',227-255,'V',257-659,'V',661-933 ##label MIS
      ##cross-references GB:M15716
GENETICS
   #gene           GDB:PGR
      ##cross-references GDB:119493
   #map_position 11q22.1-11q22.3
CLASSIFICATION     #superfamily progesterone receptor; erbA transforming protein
                     homology
KEYWORDS           alternative splicing; DNA binding; glycoprotein; nucleus;
                   phosphoprotein; steroid hormone receptor; transcription
                   regulation; zinc finger
FEATURE
   1-933               #product progesterone receptor form B #status predicted
                         #label MA1\
   165-933             #product progesterone receptor form A #status predicted
                         #label MA2\
   565-828             #domain erbA transforming protein homology #label ERBA\
   567-587             #region zinc finger CCCC motif\
   603-627             #region zinc finger CCCC motif\
   681-933             #domain steroid binding #status predicted #label STB\
   41                  #binding_site phosphate (Thr) (covalent) #status
                         predicted\
   227,232,552,733     #binding_site phosphate (Ser) (covalent) #status
                         predicted\
   329,374,601         #binding_site phosphate (Tyr) (covalent) #status
                         predicted\
   664                 #binding_site carbohydrate (Asn) (covalent) #status
                         predicted
SUMMARY            #length 933  #molecular-weight 98928  #checksum 5865

  Query Match             17.1%,  Score 666,  DB 3,  Length 933;
  Best Local Similarity  33.4%;  Pred. No. 6.63e-97;
  Matches   110,  Conservative   89,  Mismatches 113;  Indels  17;  Gaps  17;

Db    566 iclicqdeasgchygvltcgsckvffkramegqhnylcagrndcivdkirrkncpacrlr 625
          :| :|:| ||| |||| :| :|| ||||:::|:::|:|:: |:| :|| |||:| |||||
Qy    148 FCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCIIDKNRRKSCQACRLR 207

Db    626 kccqagmvl qgrkfk-kfnkvrvvraldavalpqplqvpnesqqlsqrftfspgqdiql 683
          || : |||  |:|: :  :  ||  |: |   | :  | :: | : : |        | |
Qy    208 KCYEVGMVKCGSRRERCGYRLVRRQRSADE-QL-HCAGKAKRSGGHAPRVR-ELLLDA-L 263

Db    684 ip-plinllmsiepdviyaghdntkpdtssslltslnqlgerqllsvvkwskslpqfrnl 742
           | |:  |:  ||  :  :: :: | | :|:: ||  |::::|: :: |:| :||| :|
Qy    264 SPEQLVLTLLEAEPPHVLISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVEL 322

Db    743 hiddqitllqyswmslmvfglgwrsykhvsgqmlyfapdlilneqrmk-essfyslcltm 801
           : ||: |::  || ... || |||  | .| .. |||||.|. :  |    ::  :   :
Qy    323 SLFDQVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDML 380

Db    802 wqipqefvklqvsqeeflcmkvlllntipleglrsqtq-feemrs-syirelikai-q- 857
                |  | : : |:||:| ::|||: :: | : ||  :  |  : :   :  |:
Qy    381 LATTSRFRELKLQHKEYLCVKAMILLNS-SMYPLVTATQDADSSRKLAHLLNAVTDALVW 439

Db    858 lrqkgvvsssqrfyqltklldnlhdlvkq 886
          :  |: :|| |:  :|::||  |   |::
Qy    440 VIAKSGISSQQQSMRLANLL-MLLSHVRH 467


RESULT    15
ENTRY              A25923      #type complete
TITLE              progesterone receptor - rabbit
ORGANISM           #formal_name Oryctolagus cuniculus #common_name domestic
                     rabbit
DATE               16-Aug-1988 #sequence_revision 16-Aug-1988 #text_change
                     03-Jun-1996
ACCESSIONS         A25923
REFERENCE          A25923
   #authors        Loosfelt, H.; Atger, M.; Misrahi, M.; Guiochon-Mantel, A.;
                     Meriel, C.; Logeat, F.; Benarous, R.; Milgrom, E.
   #journal        Proc. Natl. Acad. Sci. U.S.A. (1986) 83:9045-9049
   #title          Cloning and sequence analysis of rabbit progesterone receptor
                     complementary DNA.
   #cross-references MUID:87067449
   #accession      A25923
      ##molecule_type mRNA
      ##residues        1-930 ##label LOO
CLASSIFICATION     #superfamily progesterone receptor; erbA transforming protein
                     homology
KEYWORDS           DNA binding; nucleus; steroid hormone receptor; zinc finger
FEATURE
   566-826             #domain erbA transforming protein homology #label ERBA\
   568-588           #region zinc finger\
   604-628           #region zinc finger
SUMMARY          #length 930  #molecular-weight 98666  #checksum 2858

  Query Match             17.0%;  Score 662;  DB 7;  Length 930;
  Best Local Similarity  32.5%;  Pred. No. 3.74e-96;
  Matches  106;  Conservative   91,  Mismatches 114,  Indels  15,  Gaps  15;

Db     567 [illegible] 626
           :| :|:| ||  |||| :| :|| ||||:::|:::|:|:: |:| :|| |||:| |||||
Qy     148 FCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLR 207

Db     627 kccqagmvl-ggrkfk-kfnkvrvmraldavalpqpvgipnesq-ritfspsqeiq-lip 682
           || : |||  |:|: :   :  ||  |: |    | :  | :: |  :         :: | |
Qy     208 KCYEVGMVKCGSRRERCGYRLVRRQRSADE-QL-HCAGKAKRSGGHAPRVRELLLDALSP 265

Db     683 -pllnllmsiepdviyaghdntkpdtssslltslnqlgerqllsvvkwskslpgfrnlhi 741
             |:  |:  ||  :  :: :: | | :|:: ||  |::::|: :: |:| :||| :| :
Qy     266 EQLVLTLLFAEPPHVLISPPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSL 324

Db     742 ddqitliqyswmslmvfglgwrsykhvsgqmlyfapdlilneqrmk-essfyslcltmwq 800
            ||: |::  || ::: || |||  | :| :: |||||:|: :  |    ::  :   :
Qy     325 FDQVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDMLLA 382

Db     801 ipqefvklqvsqeeflcmkvllllnt-i-pleglrsqsqfeemrssyirelikaig-lrq 857
               |  | : : |:||:| ::|||: : ||      :::   : :  :  :  |:  :
Qy     383 TTSRFRELKLQHKEYLCVKAMILLNSSMYPLVIATQDADSSRKLAHLLNAVTDALVWVIA 442

Db     858 kqvvsssqrfyqltklldnlhdlvkq 883
           |: :|| |:  :|::||  |   |::
Qy     443 KSGISSQQQSMRLANLL-MLLSHVRH 467

Search completed  Thu Dec 18 09.16.32 1997
Job time : 43 secs.
```

```
************************************************************************
  _   _     ____    ____    ____    ____   _  _
 | \ / |   |  _ \  / ___|  |  _ \  / ___| | || |
 |  V  |   | |_) | \___ \  | |_) || |     | || |
 | |\/| |  |  __/   ___) | |  _ < | |___  |  __ |
 |_|  |_|  |_|     |____/  |_| \_\ \____| |_||_|  (TM)
************************************************************************

         Release 2 1D John F  Collins, Biocomputing Research Unit
         Copyright (c) 1993  1994, 1995  University of Edinburgh, U.K
                  Distribution rights by IntelliGenetics  Inc.

MPsrch_pp   protein - protein database search, using Smith-Waterman algorithm

Run on:           Thu Dec 18 09:17:41 1997;  MasPar time 5.69 Seconds
                                             451 142 Million cell updates/sec
Tabular output not generated.

Title:            >US-08-826-361A-25
Description:      (1-530) from US08826361A.pep
Perfect Score:    3904
Sequence.         1 MDIKNSPSSLNSPSSYNCSQ          ECSPAEDSKSKEGSQNPQSQ 530

Scoring table:    PAM 150
                  Gap 11

Searched:         54101 seqs, 4836506 residues

Post-processing: Minimum Match 0%
                  Listing first 45 summaries

Database:         a-issued
                    1:backl 2:51 3:52 4:53 5:54 6:55 7:56 8 PCT90 9 PCT91
                    10:PCT92 11 PCT93 12 PCT94 13:PCT95 14:PCT96

Statistics:       Mean 33 257;  Variance 151.983,  scale 0.219

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed.
      and is derived by analysis of the total score distribution.

                               SUMMARIES
                 %
Result          Query
   No.  Score  Match Length DB  ID           Description          Pred. No.
-------------------------------------------------------------------------
     1   1648   42.2    410  1  5223606-5    Patent No. 5223606    9.99e-140
     2   [illegible]   12.2    3[illegible]  1  5126[illegible]    Patent No. 5126[illegible]    2.93e-49
     3    609   15.6    375  1  5223606-3    Patent No. 5223606.   1.26e-43
     4    577   14.8    795  2  US-07-716-   Sequence 5  Applicati  3.91e-41
     5    533   13.7    532  5  US-07-952-   Sequence 2, Applicatio  5.24e-37
     6    529   13.6    446  5  US-07-952-   Sequence 4  Applicati  2.12e-36
     7    514   13.2    457  9  PCT-US91-0   Sequence 4, Applicatio  4.73e-35
     8    513   13.1    462  9  PCT-US91-0   Sequence 2  Applicati  5.81e-35
     9    486   12.4    463  9  PCT-US91-0   Sequence 6, Applicatio  1.53e-32
    10    441   11.3    454  1  5260432-2    Patent No. 5260432    1.59e-28
    11    43[illegible]   11.2    3[illegible]  1  5223606-[illegible]    Patent No. 5223606.   1.40e-28
    12    43[illegible]   11.2    448  1  5223606-2    Patent No. 5223606    1.40e-28
    13    429   11.0    462 10  PCT-US92-0   Sequence 4. Applicatio  2.28e-27
    14    428   11.0    797 10  PCT-US92-0   Sequence 2. Applicatio  2.28e-27
    15    420   10.8    440  6  US-08-333-   Sequence 8, Applicatio  1.17e-26
    16    413   10.6    462  1  5171671-2    Patent No. 5171671.   4.88e-26
    17    361    9.2    469 13  PCT-US95-1   Sequence 2, Applicatio  1.86e-21
    18    357    9.1     89  6  US-08-561-   Sequence 3, Applicatio  4.18e-21
    19    357    9.1     89  5  US-08-087-   Sequence 3, Applicatio  4.19e-21
    20    353    9.0    494 13  PCT-US95-1   Sequence 1, Applicatio  9.36e-21
    21    349    8.9     66  1  5217867-1    Patent No. 5217867.   2.09e-20
    22    329    8.4    451 13  PCT-US95-1   Sequence 2. Applicati  1.16e-19
    23    315    8.1     70  3  US-07 737-   Sequence 3, Applicatio  1.90e-17
    24    307    7.9    443  7  US-08-342-   Sequence 4, Applicatio  9.34e-17
    25    307    7.9    446 13  PCT-US95-1   Sequence 3, Applicatio  9.34e-17
    26    310    7.9    461 13  PCT-US95-1   Sequence 2, Applicatio  5.14e-17
    27    310    7.9    461 13  PCT-US95-1   Sequence 2, Applicatio  5.14e-17
    28    310    7.9    461  7  US-08-330-   Sequence 2, Applicatio  5.14e-17
    29    310    7.9    461  7  US-08-330    Sequence 2, Applicatio  5.14e-17
    30    309    7.9    548  6  US-08 333-   Sequence 2, Applicatio  6.27e-17
    31    303    7.8    410  1  5438126-2    Patent No. 5438126.   2.07e-16
    32    304    7.8    513  7  US-08-464-   Sequence 2, Applicatio  1.69e-16
    33    304    7.8    558  6  US-08-333-   Sequence 12, Applicati  1.69e-16
    34    299    7.7    460  7  US-08-342-   Sequence 2, Applicatio  4.57e-16
    35    300    7.7    601  6  US-08-333-   Sequence 14, Applicati  3.75e-16
    36    297    7.6    355  1  5223606-4    Patent No. 5223606    5.79e-16
    37    282    7.2    423 12  PCT-US94-0   Sequence 2, Applicatio  1.32e-14
    38    265    6.8    459  5  US-08 333-   Sequence 10, Applicati  3.71e-13
    39    238    6.1     89  5  US-08-087-   Sequence 2, Applicatio  7.08e-11
    40    238    6.1     89  6  US-08-561-   Sequence 2, Applicatio  7.08e-11
    41    191    4.9     60 12  PCT-US94-1   Sequence 1, Applicatio  5.41e-07
    42    169    4.3    367  3  US-07 737-   Sequence 4, Applicatio  3.17e-05
    43    141    3.6    272 13  PCT-US95-0   Sequence 23, Applicati  4.85e-03
    44    141    3.6    272  6  US-08 485    Sequence 23, Applicati  4.85e-03
    45     97    2.5    298  6  US-08-093-   Sequence 4, Applicatio  7.88e+00

                               ALIGNMENTS

RESULT    1
ID   5223606-5     STANDARD.     PRT;   444 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Patent No. 5223606.
XX
CC   Patent No. 5223606
CC        APPLICANT.  BLAUDIN DE THE, HUGHES,MARCHIO, AGNES,TIOLLAIS,
CC   PIERRE;DEJEAN, ANNE
CC        TITLE OF INVENTION: STEROID/THYROID HORMONE RECEPTOR-RELATED
CC   PROTEIN INAPPROPRIATELY EXPRESSED IN HUMAN HEPATOCELLULAR CARCINOMA
CC        NUMBER OF SEQUENCES: 11
CC        CURRENT APPLICATION DATA:
CC          APPLICATION NUMBER   US/07/134,130
CC          FILING DATE. 17-DEC-1987
CC        PRIOR APPLICATION DATA:
CC   SEQ ID NO:5:
CC          LENGTH: 410
SQ   SEQUENCE   444 AA; 50020 MW; 1110634 CN;

  Query Match             42.2%,  Score 1648,  DB 1,  Length 410;
  Best Local Similarity   56.3%,  Pred. No. 9.99e-140;
  Matches  224,  Conservative   32,  Mismatches  64,  Indels  14;  Gaps  11;

Db        1 CAVCNDYASGYHYGVWSCEGCKAFFKRSIQGHNDYMCPATNQCTIDKNRRKSCQACRLRK 60
            ||||:||| ||||||||||||||||||||||||||||:||||:||||||||||||||||||
Qy      149 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 208

Db       61 CYEVGMMKGGIRKDRRGGRMLKHKRQRDDGEGRGEVGSAGDMRAANLWPSPLMIKRSKKN 120
            |||||||:| | |::| | |:::  |||    | :    :: : :|::     :|  : |
Qy      209 CYEVGMVKCGSRRERCGYRLVR-RQRSADEQLHCAGKA-KRSGG-HAP-RV-RELL- 259

Db      121 SLALSLTADQMVSALLDAEPPILYSEYDPTRPFSEASMMGLLTNLADRELVHMINWAKRVP 180
             |  .:|...|.| .||:|||        |: ||:|||||  ||:.||:|||||||||::|
Qy      260 -ID-ALSPEQLVLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIP 317

Db      181 GFVDLTLHDQVHLLECAWLEILMIGLVWRSMEHPVKLLFAPNLLLDRNQGKCVEGMVEIF 240
            |||:|:| |||:|||  |:|:||:||:|||::|| ||:|||:|:|||::|||||::|||
Qy      318 GFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIF 377

Db      241 DMLLATSSRFRMMNLQGEEFVCLKSIILLNSGVYTFLSSTLKSLEEKDHIHRVLDKITDT 300
            ||||||.||||  .:|  |..|.:|..|||||..| .:::|  : :     : :|: :||:
Qy      378 DMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSRKLAHLLNAVTDA 436
```

```
Db      301 LIHLMAKAGLTLQQQHQRLAQLLLILSHIRHMSNKGMEHLYSMKCKNVVPLYDLLLEMLD 360
            |: ::||:|:: |||  ||| ||::|||:|| |||||||| :||||||||:||||||||:
Qy      437 LVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLN 496

Db      361 AHRLHA-PTSRGGASVEETDQSHLATAGSTSSHS 393
            || |::  :|  |:    :::|  :  || :::|
Qy      497 AHVLRGCKSSITGSECSPAEDSK SKEGSQNPQS 529


RESULT     2
ID   5223606-6      STANDARD;      PRT;   393 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Patent No. 5223606.
XX
CC   Patent No. 5223606
CC       APPLICANT:  BLAUDIN DE THE, HUGHES;MARCHIO, AGNES;TIOLLAIS,
CC   PIERRE;DEJEAN, ANNE
CC       TITLE OF INVENTION: STEROID/THYROID HORMONE RECEPTOR-RELATED
CC   PROTEIN INAPPROPRIATELY EXPRESSED IN HUMAN HEPATOCELLULAR CARCINOMA
CC       NUMBER OF SEQUENCES: 11
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/134,130
CC         FILING DATE: 17-DEC-1987
CC       PRIOR APPLICATION DATA:
CC   SEQ ID NO:6:
CC         LENGTH: 363
SQ   SEQUENCE   393 AA; 44638 MW; 907537 CN;

  Query Match             16.9%;  Score 660;  DB 1;  Length 363;
  Best Local Similarity  32.2%;  Pred. No. 2.93e-48;
  Matches   104;  Conservative   88;  Mismatches 120;  Indels   11;  Gaps   11;

Db        1 CLICGDEASGCHYGVLTCGSCKVFFKRAMEGQHNYLCAGRNDCIVDKIRRKNCPACRLRK 60
            | :|:| ||| |||| :|  :|| ||||:::|:::|:|:: |:| :|| |||:| ||||||
Qy      149 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 208

Db       61 CCQAGMVL-GGRKFK-KFNKVRVMRALDAVALPAPVGIPNESQRITFSPSQEIQLIP-PL 117
            | : |||   ::|: :  :  ||  |: |       |  :  : ::            | |  |
Qy      209 CYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 268

Db      118 INLLMSIEPDVIYAGHDNTKPDTSSSLLTSLNQLGERQLLSVVKWSKSLPGFRNLHIDDQ 177
            :  |:  ||  :  :: :: | | .|.. ||  |:...|. .. |.| .||| .| . ||
Qy      269 VLTLLEAEPPHVLISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFDQ 327

Db      178 ITLLQYSWMSLMVFGLGWRSYKHVSGQMLYFAPDLILNEQRMK-ESSFYSLCLTMWQIPQ 236
            : |::  || ::: || |||  | :| :: |||||:|: :  |    ::  :   :
Qy      328 VRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDMLLATTS 385

Db      237 EFVKLQVSQEEFLCMKVLLLLNT I PLEGLRSQSQFEEMRSSYIRELIKAIG-LRQKGV 293
             |  | : : |:|||:| ::|||: . ||       ...   .  .  .  .  |.  :  |:
Qy      386 RFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKSG 445

Db      294 VSSSQRFYQLTKLLDNLHDLVKQ 316
            :|| |:  :|::||  |   |::
Qy      446 ISSQQQSMRLANLL MLLSHVRH 467


RESULT     3
ID   5223606-7      STANDARD;      PRT;   386 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Patent No. 5223606.
XX
CC   Patent No. 5223606
CC       APPLICANT:  BLAUDIN DE THE, HUGHES;MARCHIO, AGNES;TIOLLAIS,
CC   PIERRE;DEJEAN, ANNE
CC       TITLE OF INVENTION: STEROID/THYROID HORMONE RECEPTOR-RELATED
CC   PROTEIN INAPPROPRIATELY EXPRESSED IN HUMAN HEPATOCELLULAR CARCINOMA
CC       NUMBER OF SEQUENCES: 11
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/134,130
CC         FILING DATE: 17-DEC-1987
CC       PRIOR APPLICATION DATA:
CC   SEQ ID NO:7:
CC         LENGTH: 356
SQ   SEQUENCE   386 AA; 44008 MW; 864027 CN;

  Query Match             15.6%;  Score 609;  DB 1;  Length 356;
  Best Local Similarity  36.0%;  Pred. No. 1.26e-43;
  Matches   95;  Conservative   70;  Mismatches  84;  Indels   15;  Gaps   14;

Db        1 CLVCSDEASGCHYGVLTCGSCKVFFKRAVEGQHNYLCAGRNDCIIDKIRRKNCPACRYRK 60
            | |||| ||| |||| .|  || ||||  .|...|.|:: |:| ||| |||:| ||| ||
Qy      149 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 208

Db       61 CLQAGMNL-EARKTK-K-KIKGIQQATTG-V-SQETSENPGNKT-IVPATLPQ-LTPT-L 112
            | : ||     .|. :   ..    |...   :       :  :|  :  |    | : |:|   |
Qy      209 CYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 268

Db      113 V-SLIEVIEPEVLYAGYDSSVPDSTWRIMTTLNMLGGRQVIAAVKWAKAIPGFRNLHLDD 171
            | :|||   | ||  :    ::  :::   :| :| :|: ::::   : ||| |||| :| | |
Qy      269 VLTLLEAEPPHVLISRPSAPFTEAS  MMMSLTKLADKELVHMISWAKKIPGFVELSLFD 326

Db      172 QNTLLQYSWM-FLMAALGWRSYRQSSANLLCFAPDLIIN-EQRMTLPCMYDQCKHMLYVS 229
            |  ||:  ||  || :| |||  ··· ·|· |||||::: :: : :  :  :     :|  :
Qy      327 QVRLLESCWMEVLMMGLMWRSIDHPG-KLI-FAPDLVLDRDEGKCVEGILEIFDMLLATT 384

Db      230 SELHRLQVSYEEYLCMKYLLLLSS 253
            | :: | :   ||||:| ::||:|
Qy      385 SRFRELKLQHKEYLCVKAMILLNS 408


RESULT     4
ID   US-07-716-827C-5      STANDARD;      PRT;   795 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 5, Application US/07716827C.
XX
CC    Sequence 5, Application US/07716827C
CC    Patent No. 5215916
CC     GENERAL INFORMATION:
CC       APPLICANT:  Simons Jr., Stoney S.
CC       APPLICANT:  Yamamoto, K. R.
CC       APPLICANT:  Chakraborti, P. K.
CC       APPLICANT:  Garabedian, M. J.
CC       TITLE OF INVENTION:  SUPER GLUCOCORTICOID RECEPTORS
CC       NUMBER OF SEQUENCES:  5
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Cushman, Darby & Cushman
CC         STREET:  Eleventh floor, 1615 L Street, N.W.
CC         CITY:  Washington
CC         STATE:  D.C.
CC         COUNTRY:  U.S.A.
CC         ZIP:  20036-5601
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/716,827C
CC         FILING DATE:  19910619
CC         CLASSIFICATION:  435
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Scott, Watson T.
CC          REGISTRATION NUMBER: 26,581
CC          REFERENCE/DOCKET NUMBER: WTS 5683/94453
CC        TELECOMMUNICATION INFORMATION:
CC          TELEPHONE: (202)8613000
CC          TELEFAX: (202)822-8944
CC          TELEX: 6714627 CUSH
CC      INFORMATION FOR SEQ ID NO: 5:
CC        SEQUENCE CHARACTERISTICS:
CC          LENGTH: 795 amino acids
CC          TYPE: AMINO ACID
CC          TOPOLOGY: linear
SQ   SEQUENCE   795 AA; 87498 MW; 3326890 CN;

  Query Match             14.8%;  Score 577;  DB 3;  Length 795;
  Best Local Similarity   32.9%;  Pred. No. 9.91e-41;
  Matches   92;  Conservative   83;  Mismatches   89;  Indels   16;  Gaps   15;

Db    424 SPPSSSSAATGPPPKLCLVCSDEASGCHYGVLTCGSCKVFFKRAVEGQHNYLCAGRNDCI 483
          ... .....     . ||||| ||| ||||  |  || ||||  :|   | :   : :
Qy    134 ASPVIGPGSK-PDAHPCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCT 192

Db    484 IDKIRRKNCPACRYRKCLQAGMNL-EARKTK-K-KIKGIQQATAG-V-SQDTSENPN-KT 537
          ||| |||:: ||| ||| : ||      :|: :     ::    |:.        .  :
Qy    193 IDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHA 252

Db    538 -IVPAALPQ-LTPT-LV-SLLEVIEPEVLYAGYDSSVPDSAWRIMTTLNMLGGRQVIAAV 593
            |   | . |.| || .||| | ||  .   ..  ...   :  |  |       .
Qy    253 PRVRELLLDALSPEQLVLTLLEAEPPHVLISRPSAPFTEAS--MMMSLTKLADKELVHMI 310

Db    594 KWAKAIPGLRNLHLDDQMTLLQYSWMFLMAFALGWRSYRQSSGNLLCFAPDLIIN-EQRM 652
           ||| | |: :| | ||: ||: || :: :| |||  .. |:|: ||||| .. ..
Qy    311 SWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHP-GKLI-FAPDLVLDRDEGK 368

Db    653 SLPCMYDQCKHMLFVSSELQRLQVSYEEYLCMKTLLLLSS 692
           :  :  :    :|   '    |       |||| |   || |
Qy    369 CVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNS 408

RESULT    5
ID   US-07-952-800-2     STANDARD;      PRT;   533 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 2, Application US/07952800.
XX
CC    Sequence 2, Application US/07952800
CC    Patent No. 5403925
CC     GENERAL INFORMATION:
CC       APPLICANT: OZATO, KEIKO
CC       TITLE OF INVENTION: A NEW MEMBER OF THE NUCLEAR HORMONE
CC       TITLE OF INVENTION: RECEPTOR SUPERFAMILY AND A CDNA CLONE THEREOF
CC       NUMBER OF SEQUENCES: 4
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE: TOWNSEND AND TOWNSEND
CC         STREET: 1 MARKET PLAZA, STEUART TOWER, SUITE 2000
CC         CITY: SAN FRANCISCO
CC         STATE: CA
CC         COUNTRY: USA
CC         ZIP: 94105
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE: Floppy disk
CC         COMPUTER: IBM PC compatible
CC         OPERATING SYSTEM: PC-DOS/MS-DOS
CC         SOFTWARE: PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER: US/07/952,800
CC         FILING DATE: 19920928
CC         CLASSIFICATION: 435
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME: Weber, Kenneth A
CC         REGISTRATION NUMBER: 31,677
```

```
CC          REFERENCE/DOCKET NUMBER: 15280-21-1
CC        TELECOMMUNICATION INFORMATION:
CC          TELEPHONE: 415-543-9600
CC          TELEFAX: 415-543-5043
CC      INFORMATION FOR SEQ ID NO: 2:
CC        SEQUENCE CHARACTERISTICS:
CC          LENGTH: 533 amino acids
CC          TYPE: AMINO ACID
CC          TOPOLOGY: linear
CC        MOLECULE TYPE: protein
SQ   SEQUENCE   533 AA; 56921 MW; 1406341 CN;

  Query Match             13.7%;  Score 533;  DB 5;  Length 533;
  Best Local Similarity   34.1%;  Pred. No. 9.24e-37;
  Matches  128;  Conservative   85;  Mismatches  132;  Indels   30;  Gaps   21;

Db    186 VRGLHCPPPPGGPGAGK--RLCAICGDRSSGKHYGVYSCEGCKGFFKRTIRKDLTYSCRD 243
          | | .|..|  |||. .   ::|||:|:| :|| |||| |||||||:||||:|:      |
Qy    128 VSGNPCASPVIGPGSKPDAHPCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPA 187

Db    244 NKDCIVDKRQRNRCQYCRYQKCLATGM-K----RE--AVQ-EERQRGKDKDGDGEG-AGG 294
           :.||.|| .|. || || .||    || |      ||  . :    |||: | :     | |
Qy    188 TNQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKR 247

Db    295 APEEMPVDR-ILEAELAVEQKS-DQ-GVEGPGGTGGSGSSP-NDP--VTNICQAADKQLF 348
          :    |   | :|   |: ||            | |      :  |:|  ::  : ::   ||:|
Qy    248 SGGHAPRVRELLLDALSPEQLVLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELV 307

Db    349 TLVEWAKRIPHFSSLPLDDQVILLRAGWNELLIASFSHRSIDVRDGILLATGLHVHRNSA 408
           .. ||.|| |   |.| ||| || . | |.|. ..  |||    ...|  |  . |. .
Qy    308 HMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEG 367

Db    409 HSA-GVGAIFDRVLTELVSKMRDMRMDKTELGCLRAIILFNPDAKGLSNPS-EVEV---L 463
                |. ||| .|:   |. |.....  |  |..|.||.|:     |  :: :  :     :
Qy    368 KCVEGILEIFDMLLAT-TSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKL 426

Db    464 -R--EKVYASLE-TYCKQKYP-EQQG-RFAKLLLRLPALRSIGLKCLEHLFFFKLIGDTP 517
            .  . |  .|       |    . .||. |.|.||. |: :|  : | :|||:  |    |
Qy    427 AHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVP 486

Db    518 IDTFLMEMLEAPHQL 532
          :  :|:|||:| | |
Qy    487 VYDLLLEMLNA-HVL 500

RESULT    6
ID   US-07-952-800-4     STANDARD;      PRT;   446 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 4, Application US/07952800.
XX
CC    Sequence 4, Application US/07952800
CC    Patent No. 5403925
CC     GENERAL INFORMATION:
CC       APPLICANT: OZATO, KEIKO
CC       TITLE OF INVENTION: A NEW MEMBER OF THE NUCLEAR HORMONE
CC       TITLE OF INVENTION: RECEPTOR SUPERFAMILY AND A CDNA CLONE THEREOF
CC       NUMBER OF SEQUENCES: 4
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE: TOWNSEND AND TOWNSEND
CC         STREET: 1 MARKET PLAZA, STEUART TOWER, SUITE 2000
CC         CITY: SAN FRANCISCO
CC         STATE: CA
CC         COUNTRY: USA
CC         ZIP: 94105
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE: Floppy disk
CC         COMPUTER: IBM PC compatible
CC         OPERATING SYSTEM: PC-DOS/MS-DOS
CC         SOFTWARE: PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/952,800
CC         FILING DATE:  19920928
CC         CLASSIFICATION:  435
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Weber, Kenneth A
CC         REGISTRATION NUMBER:  31,677
CC         REFERENCE/DOCKET NUMBER:  15280-21-1
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  415-543-9600
CC         TELEFAX:  415 543 5043
CC     INFORMATION FOR SEQ ID NO:  4:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  446 amino acids
CC         TYPE:  AMINO ACID
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   446 AA; 48023 MW; 964846 CN;

  Query Match             13.6%;  Score 529;  DB 5;  Length 446;
  Best Local Similarity   34.1%;  Pred. No. 2.12e-36;
  Matches   128;  Conservative   84,  Mismatches 133;  Indels  30;  Gaps  21;

Db       99 VRGLHCPPPPGGPGAGK--RLCAICGDRSSGKHYGVYSCEGCKGFFKRTIRKDLTYSCRD 156
            | |  :|::|   |||: :   ::||:|:| :|| |||| ||||||:||||:|:      | |
Qy      128 VSGNRCASFVTGPGSKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPA 187

Db      157 NKDCTVDKRQRNRCQYCRYQKCLATGM-K----RE--AVQ-EERQRGKDKDGDGDG-AGG 207
             ::||:|| :|: || || :||    || |      ||  : :    |||: | :      | |
Qy      188 TNQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKR 247

Db      208 APEEMPVDR-ILEAELAVEQKS-DQ-GVEGPGATGGGGSSP-NDP--VTNICQAADKQLF 261
            :    |  | :|   |: ||              | |     :   |:|  ::  . ..   |||.|
Qy      248 SGGHAPRVRELLLDALSPEQLVLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELV 307

Db      262 TLVEWAKRIPHFSSLPLDDQVILLRAGWNELLIASFSHRSIDVRDGILLATGLHVHRNSA 321
             :: |||:|| |  |:| ||| || : | |:|         |||         |  |  . |. .
Qy      308 HMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEG 367

Db      322 HSA-GVGAIFDRVLTELVSKMRDMRMDKTELGCLRAIILFNPDAKGL-SNPGEVEI---L 376
                |:  |||  :|:    |: |:....   |   |::|:||:|:     |  :    :  .    |
Qy      368 KCVEGILEIFDMLLAT-TSRFRELKLQHKEYLCVKAMILLNSSMYPLVIATQDADSSRKL 426

Db      377 -R--EKVYASLE-TYCKQKYP-EQQG-RFAKLLLRLPALRSIGLKCLEHLFFFKLIGDTP 430
             :   :  ' :|      |    : :||: |:|:||: |: :|   : | :|||:    |      |
Qy      427 AHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVP 486

Db      431 IDTFLMEMLEAPHQL 445
            :  :|:|||:  | |
Qy      487 VYDLLLEMLNA-HVL 500


RESULT      7
ID   PCT-US91-00399-4     STANDARD;      PRT;    467 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 4, Application PC/TUS9100399.
XX
CC    Sequence 4, Application PC/TUS9100399
CC     GENERAL INFORMATION:
CC       APPLICANT:  Mangelsdorf, Dr., David J.
CC       APPLICANT:  Evans Dr., Ronald M.
CC       TITLE OF INVENTION:  RETINOID RECEPTOR COMPOSITIONS AND
CC       TITLE OF INVENTION:  METHODS
CC       NUMBER OF SEQUENCES:  7
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  FITCH, EVEN, TABIN & FLANNERY
CC         STREET:  135 South LaSalle Street, Suite 900
CC         CITY:  Chicago
CC         STATE:  Illinois
```

```
CC         COUNTRY:  USA
CC         ZIP:  60603
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US91/00399
CC         FILING DATE:  19910122
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/478,071
CC         FILING DATE:  09-FEB-1990
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Reiter, Mr., Stephen E.
CC         REGISTRATION NUMBER:  31192
CC         REFERENCE/DOCKET NUMBER:  50852
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (619) 552-1311
CC         TELEFAX:  (619) 552 0095
CC         TELEX:  20 6566 PATLAW CGO
CC     INFORMATION FOR SEQ ID NO:  4:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  467 amino acids
CC         TYPE:  AMINO ACID
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   467 AA; 51216 MW; 1083042 CN;

  Query Match             13.2%;  Score 514;  DB 9;  Length 467;
  Best Local Similarity   33.0%;  Pred. No. 4.73e-35;
  Matches   116;  Conservative   89;  Mismatches 119,  Indels  28,  Gaps  21;

Db      138 HICAICGDRSSGKHYGVYSCEGCKGFFKRTVRKDLTYTCRDNKDCLIDKRQRNRCQYCRY 197
            |:||:|:| :|| |||| ||||||:||||:::      | |    ::| ||| :|: || ||
Qy      147 RFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRL 206

Db      198 QKCLAMGMKREAVQEER-Q-R-GK-DRNENE-VESISSAN---EDMP-VEKILEAELAVE 248
            :||  :|| : : : ||   |  : :|: :| :  :: |:        | |  :|   |: |
Qy      207 RKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPE 266

Db      249 PKTETYVEANMG--L-N-PSSP-NDP--VTNICQAADKQLFTLVEWAKRIPHFSELPLDD 301
                 | :|||:     | : ||:|  ::  : ::    |||:|  :: |||:|| | ||:| |
Qy      267 QLVLILLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 326

Db      302 QVILLRAGWNELLIASFSHRSIAVKDGILLATGLHVHRNSAHSA-GVGAIFDRVLTELVS 360
            || || : | |:|: ::  |||       :::|  | : |: :       |:  ||| :':   |
Qy      327 QVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLAT-TS 385

Db      361 KMRDMQMDKIELGCLRAIVLFNPDSKGLSNPA-EVEA---L-R--EKVYASLE-AYCKHK 412
            : |:: ::  |  |::|::|:|:      |  :: : ::    | :   : |  :|       |
Qy      386 RFRELKLQHKEYLCVKAMILLNSSMYPLVIATQDADSSRKLAHLLNAVTDALVWVIAKSG 445

Db      413 YP-EQPG-RFAKLLLRLPALRSIGLKCLEHLFFFKLIGDTPIDTFLMEMLEA 462
              : :| : |:|:||: |: :|   : | :|||:   |       |:  :|:|||:|
Qy      446 ISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNA 497


RESULT      8
ID   PCT-US91-00399-2     STANDARD;      PRT;    462 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 2, Application PC/TUS9100399.
XX
CC    Sequence 2, Application PC/TUS9100399
CC     GENERAL INFORMATION:
CC       APPLICANT:  Mangelsdorf, Dr., David J.
CC       APPLICANT:  Evans Dr., Ronald M.
CC       TITLE OF INVENTION:  RETINOID RECEPTOR COMPOSITIONS AND
CC       TITLE OF INVENTION:  METHODS
CC       NUMBER OF SEQUENCES:  7
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  FITCH, EVEN, TABIN & FLANNERY
CC         STREET:  135 South LaSalle Street, Suite 900
CC         CITY:  Chicago
CC         STATE:  Illinois
CC         COUNTRY:  USA
CC         ZIP:  60603
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US91/00399
CC         FILING DATE:  19910122
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/478,071
CC         FILING DATE:  09-FEB-1990
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Reiter, Mr., Stephen E.
CC         REGISTRATION NUMBER:  31192
CC         REFERENCE/DOCKET NUMBER:  50852
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (619) 552-1311
CC         TELEFAX:  (619) 552-0095
CC         TELEX:  20 6566 PATLAW CGO
CC     INFORMATION FOR SEQ ID NO:  2:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  462 amino acids
CC         TYPE:  AMINO ACID
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   462 AA;  50811 MW;  1063517 CN;

  Query Match             13.1%;  Score 513;  DB 9;  Length 462;
  Best Local Similarity   33.0%;  Pred. No. 5.81e-35;
  Matches  116;  Conservative   89;  Mismatches  119;  Indels   28;  Gaps   21;

Db      133 HICAICGDRSSGKHYGVYSCEGCKGFFKRTVRKDLTYTCRDNKDCLIDKRQRNRCQYCRY 192
            |:.|.:|.. . | |||| ||||||:||||::::     | |    ::| ||| :|: || ||
Qy      147 HFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRL 206

Db      193 QKCLAMGMKREAVQEER-Q-R-GK-DRNENE-VESTSSAN---EDMPVER-ILEAELAVE 243
            .||  || . .  || :  : :  |  :    |          |  : |  :  :
Qy      207 RKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPE 266

Db      244 PKTETYVEANMG--L-N-PSSP-NDP--VTNICQAADKQLFTLVEWAKRIPHFSELPLDD 296
                 | :||.    | . ||.|               |||:|  : . ||| ||: : || : :
Qy      267 QLVLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 326

Db      297 QVILLRAGWNELLIASFSHRSIAVKDGILLATGLHVHRNSAHSA-GVGAIFDRVLTELVS 355
            || ||    . .       |:|          .   .  .   .   :  |||:   :  :
Qy      327 QVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLAT-TS 385

Db      356 KMRDMQMDKTELGCLRAIVLFNPDSKGLSNPA-EVEA---L-R--EKVYASLE-AYCKHK 407
            : |.: ..  |  |..|. |:|:      :                :    |   |      :
Qy      386 RFRELKLQHKEYLCVKAMILLNSSMYPLVIATQDADSSRKLAHLLNAVTDALVWVIAKSG 445

Db      408 YP-EQPG-RFAKLLLRLPALRSIGLKCLEHLFFFKLIGDTPIDTFLMEMLEA 457
              : :| : |:|:||: |: :|  : | :|||:   |       |:   :|:||||:|
Qy      446 ISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNA 497

RESULT     9
ID   PCT-US91-00399-6     STANDARD;      PRT;   463 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
```

```
DE   Sequence 6, Application PC/TUS9100399.
XX
CC    Sequence 6, Application PC/TUS9100399
CC     GENERAL INFORMATION:
CC       APPLICANT:  Mangelsdorf, Dr., David J.
CC       APPLICANT:  Evans Dr., Ronald M.
CC       TITLE OF INVENTION:  RETINOID RECEPTOR COMPOSITIONS AND
CC       TITLE OF INVENTION:  METHODS
CC       NUMBER OF SEQUENCES:  7
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  FITCH, EVEN, TABIN & FLANNERY
CC         STREET:  135 South LaSalle Street, Suite 900
CC         CITY:  Chicago
CC         STATE:  Illinois
CC         COUNTRY:  USA
CC         ZIP:  60603
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US91/00399
CC         FILING DATE:  19910122
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/478,071
CC         FILING DATE:  09-FEB-1990
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Reiter, Mr., Stephen E.
CC         REGISTRATION NUMBER:  31192
CC         REFERENCE/DOCKET NUMBER:  50852
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (619) 552-1311
CC         TELEFAX:  (619) 552-0095
CC         TELEX:  20 6566 PATLAW CGO
CC     INFORMATION FOR SEQ ID NO:  6:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  463 amino acids
CC         TYPE:  AMINO ACID
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   463 AA;  50978 MW;  1097932 CN;

  Query Match             12.4%;  Score 486;  DB 9;  Length 463;
  Best Local Similarity   32.2%;  Pred. No. 1.53e-32;
  Matches  117;  Conservative   94;  Mismatches  120;  Indels   32;  Gaps   24;

Db      127 PSTSPGSLVK-HICAICGDRSSGKHYGVYSCEGCKGFFKRTIRKDLIYTCRDNKDCLIDK 185
            | |:|||     |:||:|:| :|| |||| ||||||:||||:|:       | |      ::| :||
Qy      136 PVTGPGSKPDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDK 195

Db      186 RQRNRCQYCRYQKCLVMGMKREAVQEER-Q-R--SRER-AESEAECASSS-HED--MP-V 236
             :|: || || :|| :||: : : : ||    |     |:| |: :   ||: : :       | |
Qy      196 NRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRV 255

Db      237 -ERILEA---E-LAVEP-KTES-YGDMNVENS--TNDPVT-NICHAADKQLFTLVEWAKR 286
             |  |  |   | |       |          . .  | .  .   .    |||:|   :: ||||:
Qy      256 RELLLDALSPEQLVLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKK 315

Db      287 IPHFSDLTLEDQVILLRAGWNELLIASFSHRSVSVQDGILLATGLHVHRSSAHS-RGVGS 345
            || |  .|.| ||| ||  .  | |.|. ..   ||:       ..:| | : |  :     |:
Qy      316 IPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILE 375

Db      346 IFDRVLTELVSKMKDMQMDKSELGCLRAIVLFNPDAKGLSNPS-EVEL---L-R--EKVY 398
            ||| :|:   |: ::: ::  |  |::|::|:|:    |  :: : ::    :  :  :  |
Qy      376 IFDMLLAT-TSRFRELKLQHKEYLCVKAMILLNSSMYPLVIATQDADSSRKLAHLLNAVT 434

Db      399 ATLE-AYTKQKYP-EQPG-RFAKLLLRLPALRSIGLKCLEHLFFFKLIGDTPIDSFLMEM 455
             .|    .|  . .| . |.|.||. |. .|  : | :|||:  |       |:  :|:||
Qy      435 DALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEM 494

Db      456 LET 458
```

```
                |::
Qy      495 LNA 497


RESULT    10
ID    5260432-2      STANDARD;      PRT;    492 AA.
XX
AC    xxxxxx
XX
DT    01-JAN-1900
XX
DE    Patent No. 5260432.
XX
CC    Patent No. 5260432
CC         APPLICANT: TAKAKU, FUMIMARO;ISHIKAWA, TAKASHI;IMAWARI, MICHIO;
CC    EVANS, RONALD M.;UMESONO, KAZUHIKO
CC         TITLE OF INVENTION: HUMAN GAMMA RETINOIC ACID RECEPTOR DNA
CC         NUMBER OF SEQUENCES: 4
CC         CURRENT APPLICATION DATA:
CC           APPLICATION NUMBER: US/07/370,407
CC           FILING DATE: 22 JUN 1989
CC    SEQ ID NO:2:
CC           LENGTH: 454
SQ    SEQUENCE   492 AA; 54527 MW; 1333563 CN;

  Query Match              11.3%;  Score 441;  DB 1;  Length 454;
  Best Local Similarity  28.4%;  Pred. No. 1.59e 28;
  Matches    99;  Conservative   88;  Mismatches 140;  Indels  22;  Gaps  18;

Db       90 CFVCNDKSSGYHYGVSSCEGCKGFFRRSIQKNMVYTCHRDKNCIINKVTRNRCQYCRLQK 149
            | ||:| :||||||| ||||||:||:|||| :   | |    : | |:|  |: || |||:|
Qy      149 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRLRK 208

Db      150 CFEVGMSKEAVRNDR N K KKKEVKEEGSPDSYELSPQLEELITKVSKAHQETF-PS-L 204
            |:|||| | : | :|    :  :::    :          : :       :|      ::: |  |
Qy      209 CYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 268

Db      205 C-QL-GKYTTNSSADHR-VQL-DLGLWDKFSELATKCIIKIVEFAKRLPGFTGLSIADQI 260
               |       :      :     : : ::    :: || | :: ::  ||::|||  ||: ||:
Qy      269 VLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFDQV 328

Db      261 TLLKAAC-LDILMLRICTRYTPEQDTMTFSDGLTLNRTQ-MHNAGFGPLTDLVFAFAGQL 318
             ||  :| :::||: :   |              : |:  | |:| : :    |:   : |:::| :::
Qy      329 RLL-ESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSRF 387

Db      319 LPLEMDDTE----TG--LL-SAICLICGDRMDLEEPEKVDKLQEPLLEAL-RLYARRRRP 370
              | ::  |        .  || |:.  .          | . . |. |        ||       |  .
Qy      388 RELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKSGIS 447

Db      371 SQP-YM-FPRMLMKITDLRGISTKGAERAITLKMEIPGPMPPLIREMLE 417
            ||    | :: :|| :: :|  | || |: : :|        |:  |: |||:
Qy      448 SQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLN 496


RESULT    11
ID    5223606-3      STANDARD;      PRT;    398 AA.
XX
AC    xxxxxx
XX
DT    01-JAN-1900
XX
DE    Patent No. 5223606.
XX
CC    Patent No. 5223606
CC         APPLICANT:  BLAUDIN DE THE, HUGHES;MARCHIO, AGNES;TIOLLAIS,
CC    PIERRE;DEJEAN, ANNE
CC         TITLE OF INVENTION: STEROID/THYROID HORMONE RECEPTOR-RELATED
CC    PROTEIN INAPPROPRIATELY EXPRESSED IN HUMAN HEPATOCELLULAR CARCINOMA
CC         NUMBER OF SEQUENCES: 11
CC         CURRENT APPLICATION DATA:
CC           APPLICATION NUMBER:  US/07/134,130
CC           FILING DATE: 17-DEC-1987
CC         PRIOR APPLICATION DATA:
CC    SEQ ID NO:3:
CC           LENGTH: 368
SQ    SEQUENCE   398 AA; 44842 MW; 907111 CN;

  Query Match              11.2%;  Score 439;  DB 1;  Length 368;
  Best Local Similarity  26.9%;  Pred. No. 2.40e 28;
  Matches    94;  Conservative   97;  Mismatches 136;  Indels  22;  Gaps  18;

Db        1 CFVCQDKSSGYHYGVSACEGCKGFFRRSIQKNMIYTCHRDKNCVINKVTRNRCQYCRLQK 60
            | || | :||||||| :|||||:||:|||| :   | |    : | |:|  |: || |||:|
Qy      149 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRLRK 208

Db       61 CFEVGMSKESVRNDR-N-K-KKKETSKQECTESYEMTAELDDLTEKIRKAHQETF-PS-L 115
            |:|||| | : | :|    :  ::: | :|          ::          : ::|     ::: |  |
Qy      209 CYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 268

Db      116 C-QL-GKYTTNSSADHR-VRL-DLGLWDKFSELATKCIIKIVEFAKRLPGFTGLTIADQI 171
               |       :      :     : : ::    :: || | :: ::  ||::|||  |:: ||:
Qy      269 VLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFDQV 328

Db      172 TLLKAAC-LDILILRICTRYTPEQDTMTFSDGLTLNRTQ-MHNAGFGPLTDLVFTFANQL 229
             ||  :| :::|:: :   |              : |:  | |:| : :    |:   : |:::: ::::
Qy      329 RLL-ESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSRF 387

Db      230 LPLEMDDTE----TG--LL-SAICLICGDRQDLEEPTKVDKLQEPLLEALK-IYIRKRRP 281
              | ::  |        :  || |::  :      || : : |:  | ::: :||  :   :   :
Qy      388 RELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKSGIS 447

Db      282 SKPH-M-FPKILMKITDLRSISAKGAERVITLKMEIPGSMPPLIQEMME 328
            |  . |  ....||  .. .|  | || |..: :|        ::  |: ||::
Qy      448 SQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLN 496


RESULT    12
ID    5223606-2      STANDARD;      PRT;    485 AA.
XX
AC    xxxxxx
XX
DT    01-JAN-1900
XX
DE    Patent No. 5223606.
XX
CC    Patent No. 5223606
CC         APPLICANT:  BLAUDIN DE THE, HUGHES;MARCHIO, AGNES;TIOLLAIS,
CC    PIERRE;DEJEAN, ANNE
CC         TITLE OF INVENTION: STEROID/THYROID HORMONE RECEPTOR-RELATED
CC    PROTEIN INAPPROPRIATELY EXPRESSED IN HUMAN HEPATOCELLULAR CARCINOMA
CC         NUMBER OF SEQUENCES: 11
CC         CURRENT APPLICATION DATA:
CC           APPLICATION NUMBER:  US/07/134,130
CC           FILING DATE: 17-DEC-1987
CC         PRIOR APPLICATION DATA:
CC    SEQ ID NO:2:
CC           LENGTH: 448
SQ    SEQUENCE   485 AA; 54437 MW; 1345912 CN;

  Query Match              11.2%;  Score 439;  DB 1;  Length 448;
  Best Local Similarity  26.9%;  Pred. No. 2.40e 28;
  Matches    94;  Conservative   97;  Mismatches 136;  Indels  22;  Gaps  18;

Db       81 CFVCQDKSSGYHYGVSACEGCKGFFRRSIQKNMIYTCHRDKNCVINKVTRNRCQYCRLQK 140
            | || | :||||||| :|||||:||:|||| :   | |    : | |:|  |: || |||:|
Qy      149 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRLRK 208

Db      141 CFEVGMSKESVRNDR-N-K-KKKETSKQECTESYEMTAELDDLTEKIRKAHQETF-PS-L 195
            |:|||| | : | :|    :  ::: | :|          ::          : ::|     ::: |  |
Qy      209 CYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 268

Db      196 C-QL-GKYTTNSSADHR-VRL-DLGLWDKFSELATKCIIKIVEFAKRLPGFTGLTIADQI 251
               |       :      :     : : ::    :: || | :: ::  ||::|||  |:: ||:
Qy      269 VLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFDQV 328

Db      252 TLLKAAC-LDILILRICTRYTPEQDTMTFSDGLTLNRTQ-MHNAGFGPLTDLVFTFANQL 309
```

```
               ||  :|  :::|:: :  |            . |.  | | |.| :  .    |.   . |...: ::::
Qy     329 RLL-ESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSRF 387

Db     310 IPIEMPETE----TG-LL-SAICLICGDRQDLEQPDKVLKLQEPLLEALK-IYIRKRRP 361
            | ::  |        :  || |::   :     || :  : |:  | ::: :||  :  :   :
Qy     388 RFLKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKSGIS 447

Db     362 SRPH-M-FPKILMKITDLRSISAKGAERVITLKMEIPGSMPPLIQEMME 408
           |  : | ::::|| :: .|  | || |:.. :|        :: |: ||..
Qy     448 SQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLN 496


RESULT   13
ID   PCT-US92-02320A-4     STANDARD.      PRT     462 AA
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 4, Application PC/TUS9202320A.
XX
CC    Sequence 4  Application PC/TUS9202320A
CC     GENERAL INFORMATION:
CC       APPLICANT:  Sloan-Kettering Institute, For Cancer Research
CC       TITLE OF INVENTION:  METHODS FOR DETECTION AND TREATMENT OF CANCER
CC       NUMBER OF SEQUENCES:  6
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  COOPER & DUNHAM
CC         STREET:  30 ROCKEFELLER PLAZA
CC         CITY:  NEW YORK
CC         STATE:  NEW YORK
CC         COUNTRY:  U.S.A.
CC         ZIP:  10112
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.24
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US92/02320A
CC         FILING DATE:
CC         CLASSIFICATION:
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 675,084
CC         FILING DATE:  22-MAR-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 673,838
CC         FILING DATE:  22-MAR-1991
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  WHITE, JOHN P
CC         REGISTRATION NUMBER:  28,678
CC         REFERENCE/DOCKET NUMBER:  38694-PCT
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 977-9550
CC         TELEFAX:  (212) 644-0525
CC         TELEX:  (212) 422523 COOP UI
CC     INFORMATION FOR SEQ ID NO:  4:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  462 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   462 AA;  50771 MW;  1101259 CN;

  Query Match            11.0%;  Score 429;  DB 10;  Length 462;
  Best Local Similarity  27.8%;  Pred. No. 2.28e-27;
  Matches   97;  Conservative   92;  Mismatches  139;  Indels   22;  Gaps   20;

Db      88 CEVCQQKSSGYHYGVSACEGCKGFFRRSIQKNMVYTCHRDKNCIINKVTRNRCQYCRLQK 147
           | || | :||||||| :|||||:||:||||  .   | |    . | | |.|   |. || ||||:|
Qy     149 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCIIDKNRRKSCQACRLRK 209

Db     148 CFEVGMSKESVRNDR-N-K-KKKEVPKPECSESYTLTPEVGPLIEKVRKAHQETF-PA-L 202
```

```
            |:||||  | : | :|    :  ::: :  |            :   |     :||     ::: |  |
Qy     209 CYEVGMVKEGSRPEPISYRLVRRLPSADLQLHDAGKAKRSGDHAPRVRELLLLALSIEIL 268

Db     199 ?-LI-GKYTINN-SFELPVSL-LILLWLKPSELSIKCIIKIVEFAKQLPGFTTLTIADQI 258
             |       :   |     :: :  :     :: |: | ::    :  ||  :||| |:: |:
Qy     269 VLTILEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFDQV 328

Db     259 TLLKAAC-LDILILRICIRYTPEQDTMIFSDGLTLNRTQ-MHNAGFGPLTDLVFAFANQL 316
            ||  :| :...|.. .  |            . |.  | | |.| .  .    |.   . |...| ::::
Qy     329 RLL-ESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSRF 387

Db     317 --LPLE-MDD--AETG-LL-SAICLICGDRQDLEQPDRVDMLQEPLLEALK-VYVRKRRP 368
             | |: ::        :  || ':: .     || :  : ::  | ::: :||  |  :  :
Qy     388 RELFLQHKEYLCVKAMILLNSSMYPLVTAIQDADSSRKLAHLLNAVTDALVWVIAKSGIS 447

Db     369 SRPH-M-FPKMLMKITDLRSISAKGAERVITLKMEIPGSMPPLIQEMLE 415
           |: . | :...|| .. :|  | || |... .|        ..  |. |||.
Qy     448 SQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLN 496


RESULT   14
ID   PCT-US92-02320A-2     STANDARD.      PRT;   797 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 2, Application PC/TUS9202320A.
XX
CC    Sequence 2, Application PC/TUS9202320A
CC     GENERAL INFORMATION:
CC       APPLICANT:  Sloan-Kettering Institute, For Cancer Research
CC       TITLE OF INVENTION:  METHODS FOR DETECTION AND TREATMENT OF CANCER
CC       NUMBER OF SEQUENCES:  6
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  COOPER & DUNHAM
CC         STREET:  30 ROCKEFELLER PLAZA
CC         CITY:  NEW YORK
CC         STATE:  NEW YORK
CC         COUNTRY:  U.S.A.
CC         ZIP:  10112
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.24
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US92/02320A
CC         FILING DATE:
CC         CLASSIFICATION:
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 675,084
CC         FILING DATE:  22-MAR-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 673,838
CC         FILING DATE:  22-MAR-1991
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  WHITE, JOHN P
CC         REGISTRATION NUMBER:  28,678
CC         REFERENCE/DOCKET NUMBER:  38694-PCT
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 977-9550
CC         TELEFAX:  (212) 644-0525
CC         TELEX:  (212) 422523 COOP UI
CC     INFORMATION FOR SEQ ID NO:  2:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  797 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   797 AA;  89281 MW;  3152972 CN;

  Query Match            11.0%;  Score 429;  DB 10;  Length 797;
  Best Local Similarity   27.8%;  Pred. No. 2.28e-27;
  Matches   97;  Conservative   92;  Mismatches  138;  Indels   22;  Gaps   20;

Db      423 CFVCQDKSSGYHYGVSACEGCKGFFRRSIQKNMVYTCHRDKNCIINKVTRNRCQYCRLQK 482
             | || |  :|||||||  :|||||:||:||||  :    | |     :  | |:|   |:  || |||:|
Qy      149 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRLRK 208

Db      483 CFEVGMSKESVRNDR-N-K-KKKEVPKPECSESYTLTPEVGELIEKVRKAHQETF-PA-L 537
            |:||||| |  :  |  :|      :    :::  :   |             :    |      :||      ::: |  |
Qy      209 CYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 268

Db      538 C-QL-GKYTTNN-SSEQRVSL-DIDLWDKFSELSTKCIIKTVEFAKQLPGFTTLTIADQI 593
                 |        :     |        :: :    :       :: |:  | ::    ·   || :|||   |:: ||:
Qy      269 VLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFDQV 328

Db      594 TLLKAAC-LDILILRICTRYTPEQDTMTFSDGLTLNRTQ-MHNAGFGPLTDLVFAFANQL 651
              ||    :|  :::|:: :   |              : |:   | |:| : :      |:   : |:::| ....
Qy      329 RLL-ESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSRF 387

Db      652 --LPLE-MDD-AETG-LL-SAICLICGDRQDLEQPDRVDMLQEPLLEALK-VYVRKRRP 703
               | |: ::        :    || |::   :       || :  :  ::    | :::  :||    |    :      :
Qy      388 RELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKSGIS 447

Db      704 SRPH-M-FPKMLMKITDLRSISAKGAERVITLKMEIPGSMPPLIQEMLE 750
            |:  :  |  ::::||  ::  :|    |  ||  |:::  :|          ··    |·  |||·
Qy      448 SQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLN 496


RESULT   15
ID   US-08-333-358-8      STANDARD;      PRT;    440 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 8, Application US/08333358.
XX
CC    Sequence 8, Application US/08333358
CC    Patent No. 5571696
CC     GENERAL INFORMATION:
CC       APPLICANT:  EVANS Ph.D., RONALD M.
CC       APPLICANT:  MANGELSDORF Ph.D., DAVID J.
CC       APPLICANT:  ONG Ms., ESTELITA S.
CC       APPLICANT:  ORO Ph.D., ANTHONY E.
CC       APPLICANT:  BORGMEYER Ph.D., UWE K.
CC       APPLICANT:  GIGUERE Ph.D., VINCENT NMN
CC       APPLICANT:  YAO Mr., TSO-PANG NMN
CC       TITLE OF INVENTION:  NOVEL RECEPTORS
CC       NUMBER OF SEQUENCES:  14
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Pretty, Schroeder, Brueggemann & Clark
CC         STREET:  444 So. Flower St., Suite 2000
CC         CITY:  Los Angeles
CC         STATE:  CA
CC         COUNTRY:  US
CC         ZIP:  90071-2921
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/333,358
CC         FILING DATE:
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/761,068
CC         FILING DATE:
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Reiter Ph.D., Stephen E.
CC         REGISTRATION NUMBER:  31192
CC         REFERENCE/DOCKET NUMBER:  P31 8936
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (619) 535-9001
CC         TELEFAX:  (619) 535-8949
CC     INFORMATION FOR SEQ ID NO:  8:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  440 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   440 AA;  49883 MW;  957063 CN;

  Query Match             10.8%;  Score 420;  DB 6;  Length 440;
  Best Local Similarity   30.9%;  Pred. No. 1.17e-26;
  Matches   84;  Conservative   68;  Mismatches  104;  Indels   17;  Gaps   15;

Db       80 GPAPKMIGNELCSVCGDKASGFHYNVLSCEGCKGFFRRSVIKGAHYICHSGGHCPMDTYM 139
             ||::|        :|:|||:| |||:|| | ||||||:||:||:           ||| :    :| :!
Qy      139 GEGSRE-LAHFTAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNR 197

Db      140 RRKCQECRLRKCRQAGMREECVLSEEQ--IRLKKLKRQEEEQAHATSLPPRRSSPPQILP 197
            |: || |||||| : ||      |        |:      || :   |    :|| |  ::  :  |  :: :   :
Qy      198 RKSCQACRLRKCYEVGM-VKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVR 256

Db      198 QLSPEQLGMIEKLVAAQQQCNRRSFSDRLRVTPWPMAPDPHSREARQQRFAHFTELAIVS 257
            :|   : |:   | || :  :  :   : :    | :: :| :: | : ||          : :| ||
Qy      257 ELLLDALSP-EQLVLTLLEAE--P-PHVL-IS-RP-SA-PFT-EASM-MMS-LTKLADKE 305

Db      258 VQEIVDFAKQLPGFLQLSREDQIALLKTSAIEVMLLETSRRYNPGSESITFLKDFSYNRE 317
            :   ::    || :|||::||    ||: || :   :||:::     :|        :   :  |   |:  :|:
Qy      306 LVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRD 365

Db      318 DFAKAGLQV-EFINPIFEFSRAMNELQLNDAEF 349
            : :|      : |::: :: :     || |    |:
Qy      366 E-GKCVEGILEIFDMLLATTSRFRELKLQHKEY 397

Search completed: Thu Dec 18 09.18.01 1997
Job time : 20 secs.
```

```
*****************************************************************
                       MPSRCH                          (TM)
*****************************************************************

         Release 2.1D John F. Collins, Biocomputing Research Unit
         Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
                 Distribution rights by IntelliGenetics, Inc.

MPsrch_pp   protein - protein database search, using Smith-Waterman algorithm

Run on:          Thu Dec 18 08:29:09 1997;  MasPar time 3.43 Seconds
                                            406.313 Million cell updates/sec
Tabular output not generated

Title:           >US-08-826-361A-3
Description:     (1-66) from US08826361A.pep
Perfect Score:   535
Sequence:        1 CAVCSDYASGYHYGVWSCEG.........KNRRKSCQACRLRKCYEVGM 66

Scoring table:   PAM 150
                 Gap 11

Searched:        59021 seqs, 21210388 residues

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database:        swiss-prot34
                   1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
                   8:part8 9:part9 10:part10 11:part11

Statistics:      Mean 35.216;  Variance 56.191;  scale 0.627

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                              SUMMARIES
                  %
Result           Query
  No.   Score    Match Length DB  ID           Description              Pred. No.
-----------------------------------------------------------------------------------
    1    527     98.5    575   3  ESTR_ORYLA   ESTROGEN RECEPTOR (ER  4.20e-106
    2    527     98.5    589   3  ESTR_CHICK   ESTROGEN RECEPTOR (ER  4.20e-106
    3    527     98.5    595   3  ESTR_HUMAN   ESTROGEN RECEPTOR (ER  4.20e-106
    4    527     98.5    599   3  ESTR_MOUSE   ESTROGEN RECEPTOR (ER  4.20e-106
    5    527     98.5    600   3  ESTR_RAT     ESTROGEN RECEPTOR (ER  4.20e-106
    6    517     96.6    574   3  ESTR_SALIR   ESTROGEN RECEPTOR (ER  1.48e-103
    7    507     94.8    535   3  ESTR_SALSA   ESTROGEN RECEPTOR (ER  5.19e-101
    8    496     92.7    111   3  ESTR_SHEEP   ESTROGEN RECEPTOR (ER  2.23e-98
    9    387     72.3    433   3  ERR2_HUMAN   STEROID HORMONE RECEP  5.16e-71
   10    374     69.9    521   3  ERR1_HUMAN   STEROID HORMONE RECEP  1.39e-67
   11    352     65.8    786   7  PRGR_CHICK   PROGESTERONE RECEPTOR  3.81e-62
   12    352     65.8    923   7  PRGR_MOUSE   PROGESTERONE RECEPTOR  3.81e-62
   13    352     65.8    930   7  PRGR_RABIT   PROGESTERONE RECEPTOR  3.81e-62
   14    352     65.8    933   7  PRGR_HUMAN   PROGESTERONE RECEPTOR  3.81e-62
   15    351     65.6    709   1  ANDR_RABIT   ANDROGEN RECEPTOR (FR  6.71e-62
   16    351     65.6    899   1  ANDR_MOUSE   ANDROGEN RECEPTOR.     6.71e-62
   17    351     65.6    902   1  ANDR_RAT     ANDROGEN RECEPTOR      6.71e-62
   18    351     65.6    919   1  ANDR_HUMAN   ANDROGEN RECEPTOR      6.71e-62
   19    349     65.2    776   4  GCR_XENLA    GLUCOCORTICOID RECEPT  2.09e-61
   20    349     65.2    777   4  GCR_HUMAN    GLUCOCORTICOID RECEPT  2.09e-61
   21    349     65.2    795   4  GCR_RAT      GLUCOCORTICOID RECEPT  2.09e-61
   22    343     64.1    981   6  MCR_RAT      MINERALOCORTICOID REC  6.23e-60
   23    343     64.1    984   6  MCR_HUMAN    MINERALOCORTICOID REC  6.23e-60
   24    339     63.4    783   4  GCR_MOUSE    GLUCOCORTICOID RECEPT  5.98e-59
   25    338     63.2    443   9  RRG2_HUMAN   RETINOIC ACID RECEPTO  1.05e-58
   26    338     63.2    447   9  RRG2_MOUSE   RETINOIC ACID RECEPTO  1.05e-58
   27    338     63.2    454   9  RRG1_HUMAN   RETINOIC ACID RECEPTO  1.05e-58
   28    338     63.2    454   9  RRG1_MOUSE   RETINOIC ACID RECEPTO  1.05e-58
   29    336     62.8    441   9  RRG2_NOTVI   RETINOIC ACID RECEPTO  3.25e-58
   30    336     62.8    505   9  RRG1_NOTVI   RETINOIC ACID RECEPTO  3.25e-58
   31    336     62.8    598   6  NGFI_CANFA   ORPHAN NUCLEAR RECEPT  3.25e-58
   32    335     62.6    598   6  NGFI_HUMAN   EARLY RESPONSE PROTEI  5.25e-58
   33    334     62.4    771   4  GCR_CAVPO    GLUCOCORTICOID RECEPT  1.01e-57
   34    333     62.2    442   8  RRG2_XENLA   RETINOIC ACID RECEPTO  1.77e-57
   35    333     62.2    448   8  RRB2_HUMAN   RETINOIC ACID RECEPTO  1.77e-57
   36    333     62.2    449   8  RRB2_CHICK   RETINOIC ACID RECEPTO  1.77e-57
   37    333     62.2    455   8  RRB1_CHICK   RETINOIC ACID RECEPTO  1.77e-57
   38    333     62.2    458   9  RRA_XENLA    RETINOIC ACID RECEPTO  1.77e-57
   39    333     62.2    462   9  RRA1_HUMAN   RETINOIC ACID RECEPTO  1.77e-57
   40    333     62.2    476   8  RRG1_XENLA   RETINOIC ACID RECEPTO  1.77e-57
   41    333     62.2    482   9  RRB_MOUSE    RETINOIC ACID RECEPTO  1.77e-57
   42    333     62.2    577   6  NGFI_XENLA   NERVE GROWTH FACTOR I  1.77e-57
   43    332     62.1    462   8  RRA_MOUSE    RETINOIC ACID RECEPTO  3.10e-57
   44    332     62.1    597   6  NGFI_RAT     NERVE GROWTH FACTOR I  3.10e-57
   45    332     62.1    601   6  NGFI_MOUSE   NUCLEAR HORMONE RECEP  3.10e-57

                              ALIGNMENTS

RESULT    1
ID   ESTR_ORYLA     STANDARD;      PRT;   575 AA.
AC   P50241;
DT   01-OCT-1996 (REL. 34, CREATED)
DT   01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   ESTROGEN RECEPTOR (ER)
GN   ESR.
OS   ORYZIAS LATIPES (MEDAKA FISH).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; PISCES; GNATHOSTOMATA;
OC   OSTEICHTHYES; ACTINOPTERYGII; ATHERINIFORMES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RA   OKADA H.;
RL   SUBMITTED (APR-1994) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; D28954; G602920; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
FT   DOMAIN        1     140       MODULATING.
FT   DNA_BIND    141     206       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     141     161       C4-TYPE.
FT   ZN_FING     177     201       C4-TYPE.
FT   DOMAIN      207     575       STEROID-BINDING.
FT   DOMAIN      254     262       POLY-GLY.
FT   DOMAIN      275     278       POLY-LEU.
FT   DOMAIN      466     470       POLY-LEU.
FT   DOMAIN      531     534       POLY-GLY.
SQ   SEQUENCE   575 AA;  62487 MW;  D4FE41D1 CRC32;

  Query Match             98.5%;  Score 527;  DB 3;  Length 575;
  Best Local Similarity   97.0%;  Pred. No. 4.20e-106;
  Matches   64;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Db        141 cavcsdyasgyhygvwscegckaffkrsiqghndymcpatnqctidrnrrkscqacrlrk 200
              |||||||||||||||||||||||||||||||||||:||||||||||:|||||||||||||
Qy          1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 60
```

```
CC       COMPLEX APPEARS TO RECOGNIZE DISCRETE DNA SEQUENCES UPSTREAM OF
CC       TRANSCRIPTIONAL START SITES.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS
DR   EMBL; X03805; G63380; -
DR   PIR, S07192, S07192.
DR   HSSP; P06536; 1GDC.
DR   PROSITE; PS00031; STEROID_FINGER
DR   TRANSFAC; T00264; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC FINGER, STEROID-BINDING.
FT   DOMAIN        1     178       MODULATING.
FT   DNA_BIND    179     244       C4 TYPE ZINC FINGERS (TWO).
FT   ZN_FING     179     199       C4 TYPE.
FT   ZN_FING     215     239       C4-TYPE.
FT   DOMAIN      245     304       HINGE.
FT   DOMAIN      305     589       STEROID-BINDING.
SQ   SEQUENCE   589 AA;  66746 MW;  85FF1F4A CRC32;

  Query Match             98.5%;  Score 527;  DB 3;  Length 589;
  Best Local Similarity   97.0%;  Pred. No. 4.20e-106;
  Matches   64;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Db     179 cavcndyasgyhygvwscegckaffkrsiqghndymcpatnqctidknrrkscqacrlrk 238
           ||||:||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Qy       1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 60

Db     239 cyevgm 244
           ||||||
Qy      61 CYEVGM 66


RESULT    3
ID   ESTR_HUMAN     STANDARD;      PRT;   595 AA.
AC   P03372;
DT   21-JUL-1986 (REL. 01, CREATED)
DT   01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DT   01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE   ESTROGEN RECEPTOR (ER).
GN   ESR.
```

```
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- IN THE ABSENCE OF LIGAND, STEROID HORMONE RECEPTORS ARE THOUGHT
CC       TO BE WEAKLY ASSOCIATED WITH NUCLEAR COMPONENTS; HORMONE BINDING
CC       GREATLY INCREASES RECEPTOR AFFINITY. THE HORMONE-RECEPTOR
CC       COMPLEX APPEARS TO RECOGNIZE DISCRETE DNA SEQUENCES UPSTREAM OF
CC       TRANSCRIPTIONAL START SITES.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; X03635; G31234; -.
DR   EMBL, M12674, G182193; -.
DR   PIR, A03244; QRHUE.
DR   HSSP; P06536; 1GDC.
DR   MIM; 133430; -.
DR   PROSITE; PS00031; STEROID_FINGER.
DR   TRANSFAC; T00261; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER, STEROID-BINDING; POLYMORPHISM.
FT   DOMAIN        1     184       MODULATING.
FT   DNA_BIND    185     250       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     185     205       C4-TYPE.
FT   ZN_FING     221     245       C4-TYPE.
FT   DOMAIN      251     310       HINGE.
FT   DOMAIN      311     551       STEROID-BINDING.
FT   VARIANT     400     400       G -> V (DESTABILIZES THE RECEPTOR AND
FT                                 DECREASES ITS AFFINITY FOR ESTRADIOL AT
FT                                 25 DEGREES CELSIUS, BUT NOT AT 4 DEGREES
FT                                 CELSIUS).
FT   VARIANT     447     447       C -> A (LOSS OF HORMONE BINDING CAPACITY
FT                                 AND TEMPERATURE-SENSITIVE LOSS IN DNA-
FT                                 BINDING).
SQ   SEQUENCE   595 AA;  66216 MW;  D6945BB9 CRC32;

  Query Match             98.5%;  Score 527;  DB 3;  Length 595;
  Best Local Similarity   97.0%;  Pred. No. 4.20e-106;
  Matches   64;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Db     185 cavcndyasgyhygvwscegckaffkrsiqghndymcpatnqctidknrrkscqacrlrk 244
```

```
           ||||:||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Qy       1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRLRK 60

Db     245 cyevgm 250
           ||||||
Qy      61 CYEVGM 66


RESULT     4
ID   ESTR_MOUSE     STANDARD;      PRT;   599 AA.
AC   P19785;
DT   01-FEB-1991 (REL. 17, CREATED)
DT   01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
DT   01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE   ESTROGEN RECEPTOR (ER)
GN   ESR OR ESTR.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=UTERUS;
RX   MEDLINE; 91042558.
RA   WHITE R., LEES J.A., NEEDHAM M., HAM J., PARKER M.;
RL   MOL. ENDOCRINOL. 1:735-744(1987).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- IN THE ABSENCE OF LIGAND, STEROID HORMONE RECEPTORS ARE THOUGHT
CC       TO BE WEAKLY ASSOCIATED WITH NUCLEAR COMPONENTS. HORMONE BINDING
CC       GREATLY INCREASES RECEPTOR AFFINITY. THE HORMONE-RECEPTOR
CC       COMPLEX APPEARS TO RECOGNIZE DISCRETE DNA SEQUENCES UPSTREAM OF
CC       TRANSCRIPTIONAL START SITES
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS
DR   EMBL; M38651; G193180; -.
DR   PIR; A40061; A40061.
DR   HSSP; P06536; 1GDC.
DR   PROSITE; PS00031; STEROID_FINGER.
DR   TRANSFAC; T00259; -
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING
FT   DOMAIN          1    188       MODULATING.
FT   DNA_BIND      189    254       C4 TYPE ZINC FINGERS (TWO).
FT   ZN_FING       189    209       C4-TYPE.
FT   ZN_FING       225    249       C4-TYPE
FT   DOMAIN        255    599       STEROID-BINDING.
FT   DOMAIN         64     72       POLY-ALA.
SQ   SEQUENCE   599 AA;  66955 MW;  6E19FAEB CRC32;

  Query Match             98.5%;  Score 527;  DB 3;  Length 599;
  Best Local Similarity   97.0%;  Pred. No. 4.20e-100;
  Matches   64;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Db     189 cavcndyasgyhygvwscegckaffkrsiqghndymcpatnqctidknrrkscqacrlrk 248
           ||||:||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Qy       1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRLRK 60

Db     249 cyevgm 254
           ||||||
Qy      61 CYEVGM 66


RESULT     5
ID   ESTR_RAT       STANDARD;      PRT;   600 AA.
AC   P06211;
DT   01-JAN-1988 (REL. 06, CREATED)
DT   01-JAN-1988 (REL. 06, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   ESTROGEN RECEPTOR (ER).
GN   ESR OR ESTR.
OS   RATTUS NORVEGICUS (RAT).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WISTAR;
RA   MURAMATSU M.;
RL   SUBMITTED (MAR-1987) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 87174780.
RA   KOIKE S., SAKAI M.;
RL   NUCLEIC ACIDS RES. 15:2499-2513(1987).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=UTERUS;
RA   MAGGI A.M.A.;
RL   SUBMITTED (JUN-1991) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- IN THE ABSENCE OF LIGAND, STEROID HORMONE RECEPTORS ARE THOUGHT
CC       TO BE WEAKLY ASSOCIATED WITH NUCLEAR COMPONENTS. HORMONE BINDING
CC       GREATLY INCREASES RECEPTOR AFFINITY. THE HORMONE-RECEPTOR
CC       COMPLEX APPEARS TO RECOGNIZE DISCRETE DNA SEQUENCES UPSTREAM OF
CC       TRANSCRIPTIONAL START SITES.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; Y00102; G56111; -.
DR   EMBL; X61098; G56121; -.
DR   PIR; S07379; S07379.
DR   HSSP; P06536; 1GDC.
DR   PROSITE; PS00031; STEROID_FINGER.
DR   TRANSFAC; T00258; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING
FT   DOMAIN          1    189       MODULATING.
FT   DNA_BIND      190    255       C4 TYPE ZINC FINGERS (TWO).
FT   ZN_FING       190    210       C4-TYPE.
FT   ZN_FING       226    250       C4-TYPE.
FT   DOMAIN        256    600       STEROID-BINDING.
FT   DOMAIN         64     71       POLY-ALA.
FT   CONFLICT      488    488       N -> T (IN REF. 3).
SQ   SEQUENCE   600 AA;  67030 MW;  4C0C59B2 CRC32;

  Query Match             98.5%;  Score 527;  DB 3;  Length 600;
  Best Local Similarity   97.0%;  Pred. No. 4.20e-100;
  Matches   64;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Db     190 cavcndyasgyhygvwscegckaffkrsiqghndymcpatnqctidknrrkscqacrlrk 249
           ||||.||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Qy       1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRLRK 60

Db     250 cyevgm 255
           ||||||
Qy      61 CYEVGM 66


RESULT     6
ID   ESTR_SALIR     STANDARD;      PRT;   574 AA.
AC   P16058;
DT   01-APR-1990 (REL. 14, CREATED)
DT   01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   ESTROGEN RECEPTOR (ER).
GN   ESR.
OS   SALMO IRIDEUS (RAINBOW TROUT).
```

```
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; PISCES; GNATHOSTOMATA;
OC   OSTEICHTHYES; ACTINOPTERYGII; SALMONIFORMES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 91006824.
RA   PAKDEL F., LE GAC F., LE GOFF P., VALOTAIRE Y.;
RL   MOL. CELL. ENDOCRINOL. 71:195-204(1990).
RN   [2]
RP   SEQUENCE OF 150-574 FROM N.A.
RC   TISSUE=LIVER;
RX   MEDLINE; 89127284.
RA   PAKDEL F., LE GUELLEC C., VAILLANT C., LE ROUX M.G., VALOTAIRE Y.;
RL   MOL. ENDOCRINOL. 3:44-51(1989).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- IN THE ABSENCE OF LIGAND, STEROID HORMONE RECEPTORS ARE THOUGHT
CC       TO BE WEAKLY ASSOCIATED WITH NUCLEAR COMPONENTS; HORMONE BINDING
CC       GREATLY INCREASES RECEPTOR AFFINITY. THE HORMONE-RECEPTOR
CC       COMPLEX APPEARS TO RECOGNIZE DISCRETE DNA SEQUENCES UPSTREAM OF
CC       TRANSCRIPTIONAL START SITES.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; M31559; G213784; -.
DR   PIR; A37197; A37197.
DR   PIR; A40070; A40070.
DR   HSSP; P06536; 1GDC.
DR   PROSITE; PS00031; STEROID_FINGER.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
FT   DOMAIN          1    146       MODULATING (POTENTIAL).
FT   DNA_BIND      147    212       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING       147    167       C4-TYPE.
FT   ZN_FING       183    207       C4-TYPE.
FT   DOMAIN        213    574       STEROID-BINDING (POTENTIAL).
SQ   SEQUENCE   574 AA;  62872 MW;  B391AD02 CRC32;

  Query Match             96.6%;  Score 517;  DB 3;  Length 574;
  Best Local Similarity   93.9%;  Pred. No. 1.48e-103;
  Matches   62;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Db    147 cavcsdfasgyhygvwscegckaffkrsiqghndymcpatnqctmdrnrrkscqacrlrk 206
          ||||||:||||||||||||||||||||||||||||:||||||||:|:|||||||||||||
Qy      1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRLRK 60

Db    207 cyevgm 212
          ||||||
Qy     61 CYEVGM 66


RESULT    7
ID   ESTR_SALSA     STANDARD;      PRT;   535 AA.
AC   P50242;
DT   01-OCT-1996 (REL. 34, CREATED)
DT   01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   ESTROGEN RECEPTOR (ER) (FRAGMENT).
GN   ESR.
OS   SALMO SALAR (ATLANTIC SALMON).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; PISCES; GNATHOSTOMATA;
OC   OSTEICHTHYES; ACTINOPTERYGII; SALMONIFORMES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RA   ROGERS S.A., LLEWELLYN L., RAMSURN V.P., SWEENEY G.E., WIGHAM T.;
RL   SUBMITTED (JUL-1995) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; X89959; G929902; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
FT   NON_TER         1      1
FT   DOMAIN         <1    104       MODULATING (POTENTIAL).
FT   DNA_BIND      105    170       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING       105    125       C4-TYPE.
FT   ZN_FING       141    165       C4-TYPE.
FT   DOMAIN        171    535       STEROID-BINDING (POTENTIAL).
SQ   SEQUENCE   535 AA;  58900 MW;  0E71C0DC CRC32;

  Query Match             94.8%;  Score 507;  DB 3;  Length 535;
  Best Local Similarity   92.4%;  Pred. No. 5.19e-101;
  Matches   61;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Db    105 cavcsdfasgyhygfwscegckaffkrsiqghndymcpatnqctmdrnrrkscqacrlrk 164
          ||||||:||||||| ||||||||||||||||||||:||||||||:|:|||||||||||||
Qy      1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRLRK 60

Db    165 cyevgm 170
          ||||||
Qy     61 CYEVGM 66


RESULT    8
ID   ESTR_SHEEP     STANDARD;      PRT;   111 AA.
AC   P49885;
DT   01-OCT-1996 (REL. 34, CREATED)
DT   01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   ESTROGEN RECEPTOR (ER) (FRAGMENT).
GN   ESR.
OS   OVIS ARIES (SHEEP).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; ARTIODACTYLA.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   ING N.H., SPENCER T.E., BAZER F.W.;
RL   SUBMITTED (JUN-1995) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; U30299; G927280; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
FT   NON_TER         1      1
FT   DOMAIN         <1     49       MODULATING.
FT   DNA_BIND       50    111       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING        50     70       C4-TYPE.
FT   ZN_FING        86    110       C4-TYPE.
FT   NON_TER       111    111
SQ   SEQUENCE   111 AA;  12527 MW;  2D129579 CRC32;

  Query Match             92.7%;  Score 496;  DB 3;  Length 111;
  Best Local Similarity   96.8%;  Pred. No. 3.23e-98;
  Matches   60;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Db     50 cavcndyasgyhygvwscegckaffkrsiqghndymcpatnqctidknrrkscqacrlrk 109
          ||||:||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Qy      1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRLRK 60
```

```
Db    110 cy 111
          ||
Qy     61 CY 62


RESULT    9
ID   ERR2_HUMAN     STANDARD;      PRT;   433 AA.
AC   P11475;
DT   01-OCT-1989 (REL. 12, CREATED)
DT   01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)
DT   01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE   STEROID HORMONE RECEPTOR ERR2
GN   ERR2.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=HEART;
RX   MEDLINE; 88122546.
RA   GIGUERE V., YANG N., SEGUI P., EVANS R.M.;
RL   NATURE 331:91-94(1988).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; X51417; G36611; -.
DR   PIR; B29345; B29345.
DR   HSSP; P06536; 1GDC.
DR   PROSITE; PS00031; STEROID_FINGER.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER.
FT   DNA_BIND    103    168       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     103    123       C4-TYPE.
FT   ZN_FING     139    163       C4-TYPE.
SQ   SEQUENCE   433 AA;  48287 MW;  673BF88F CRC32;

  Query Match             72.3%;  Score 387;  DB 3;  Length 433;
  Best Local Similarity  72.7%;  Pred. No. 8.26e-71;
  Matches    48;  Conservative    7;  Mismatches  11;  Indels    0;  Gaps    0;

Db    103 clvcgdiasgyhygvasceackaffkrtiqgnieyscpatnceitkrrrksсqacrfmk 162
          | ||:| |||||||| |||:|||||||||:|||: :| |||||:| | | |||||||||: |
Qy      1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 60

Db    163 clkvgm 168
          |  |||
Qy     61 CYEVGM 66


RESULT   10
ID   ERR1_HUMAN     STANDARD;      PRT;   521 AA.
AC   P11474;
DT   01-OCT-1989 (REL. 12, CREATED)
DT   01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)
DT   01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE   STEROID HORMONE RECEPTOR ERR1.
GN   ERR1.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=KIDNEY;
RX   MEDLINE; 88122546.
RA   GIGUERE V., YANG N., SEGUI P., EVANS R.M.;
RL   NATURE 331:91-94(1988).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; X51416; G36609; -.
DR   PIR; A29345; A29345.
DR   HSSP; P06536; 1GDC.
DR   PROSITE; PS00031; STEROID_FINGER.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER.
FT   DNA_BIND    176    241       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     176    196       C4-TYPE.
FT   ZN_FING     212    236       C4-TYPE.
SQ   SEQUENCE   521 AA;  55742 MW;  E2D723EA CRC32;

  Query Match             69.9%;  Score 374;  DB 3;  Length 521;
  Best Local Similarity  69.7%;  Pred. No. 1.39e-67;
  Matches    46;  Conservative    8;  Mismatches  12;  Indels    0;  Gaps    0;

Db    176 clvcgdvasgyhygvasceackaffkrtiqgsieyscpasnceitkrrrkacqacrftk 235
          | ||:| |||||||| |||:||||||||:|||  :| |||:|:| | | |||:||||: |
Qy      1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 60

Db    236 clrvgm 241
          |  |||
Qy     61 CYEVGM 66


RESULT   11
ID   PRGR_CHICK     STANDARD;      PRT;   786 AA.
AC   P07812;
DT   01-AUG-1988 (REL. 08, CREATED)
DT   01-AUG-1988 (REL. 08, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   PROGESTERONE RECEPTOR (PR).
GN   PGR.
OS   GALLUS GALLUS (CHICKEN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;
OC   GALLIFORMES.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   GRONEMEYER H., TURCOTTE B., QUIRIN-STRICKER C., BOCQUEL M.T.;
RL   NUCLEIC ACIDS RES. 13:3985-3994(1985).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 88166640.
RA   GRONEMEYER H., TURCOTTE B., QUIRIN-STRICKER C., BOCQUEL M.T.,
RA   MEYER M.E., KROZOWSKI Z., JELTSCH J.M., LEROUGE T., GARNIER J.M.,
RA   CHAMBON P.;
RL   EMBO J. 6:3985-3994(1987).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 91042592.
RA   CONNEELY O.M., DOBSON A.D.W., TSAI M.-J., BEATTIE W.G., TOFT D.O.,
RA   HUCKABY C.S., ZARUCKI T., SCHRADER W.T., O'MALLEY B.W.;
RL   MOL. ENDOCRINOL. 1:517-525(1987).
RN   [4]
RP   SEQUENCE OF 128-164 FROM N.A.
RX   MEDLINE; 86289413.
RA   CONNEELY O.M., SULLIVAN W.P., TOFT D.O., BIRNBAUMER M., COOK R.G.,
RA   MAXWELL B.L., ZARUCKI-SCHULZ T., GREENE G.L., SCHRADER W.T.,
RA   O'MALLEY B.W.;
RL   SCIENCE 233:767-770(1986).
RN   [5]
RP   SEQUENCE OF 417-490 FROM N.A.
RX   MEDLINE; 86287271.
RA   JELTSCH J.M., KROZOWSKI Z., QUIRIN-STRICKER C., GRONEMEYER H.,
RA   SIMPSON R.J., GARNIER J.M., KRUST A., JACOB F., CHAMBON P.;
RL   PROC. NATL. ACAD. SCI. U.S.A. 83:5424-5428(1986).
RN   [6]
RP   DIFFERENCE BETWEEN FORM 1 AND FORM 2.
RX   MEDLINE; 89340509.
RA   CONNEELY O.M., KETTELBERGER D.M., TSAI M.-J., SCHRADER W.T.,
RA   O'MALLEY B.W.;
RL   J. BIOL. CHEM. 264:14062-14064(1989).
RN   [7]
RP   SEQUENCE FROM N.A. (ALL FORMS).
RX   MEDLINE; 90154085.
RA   JELTSCH J.-M., TURCOTTE B., GARNIER J.-M., LEROUGE T., KROZOWSKI Z.,
RA   GRONEMEYER H., CHAMBON P.;
RL   J. BIOL. CHEM. 265:3967-3974(1990).
```

```
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN
CC   -!- ALTERNATIVE PRODUCTS: THERE ARE TWO FORMS OF THE PROGESTERONE
CC       RECEPTOR: FORM A (THAT STARTS WITH MET-1) AND FORM B (THAT STARTS
CC       WITH MET-128). THE SEQUENCE SHOWN IS THAT OF FORM A
CC   -!- TISSUE SPECIFICITY: OVIDUCT.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; Y00092; G63745; -.
DR   EMBL; M13972; G212598; -.
DR   EMBL; M37518; G212561; -.
DR   EMBL; M14278; G212600; -.
DR   EMBL; M14279; G212606; -.
DR   EMBL; M14280; G212608; -.
DR   EMBL; M32732; G212558; -.
DR   EMBL; M31104; G212558; JOINED.
DR   EMBL; M32726; G212558; JOINED.
DR   EMBL; M32727; G212558; JOINED.
DR   EMBL; M32728; G212558; JOINED.
DR   EMBL; M32729; G212558; JOINED.
DR   EMBL; M32730; G212558; JOINED.
DR   EMBL; M32732; G212559; -.
DR   EMBL; M31104; G212559; JOINED.
DR   EMBL; M32726; G212559; JOINED.
DR   EMBL; M32727; G212559; JOINED.
DR   EMBL; M32728; G212559; JOINED.
DR   EMBL; M32729; G212559; JOINED.
DR   EMBL; M32730; G212559; JOINED.
DR   EMBL; M31104; G212556; -.
DR   EMBL; M31104; G212557; -.
DR   PIR; A24312; A24312.
DR   PIR; A24661; A24661.
DR   PIR; S06284; S06284.
DR   HSSP; P06536; 1GDC.
DR   PROSITE; PS00031; STEROID_FINGER.
DR   TRANSFAC; T00698; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING; ALTERNATIVE SPLICING.
FT   DOMAIN        1    420       MODULATING, PRO-RICH.
FT   DNA_BIND    421    486       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     421    441       C4-TYPE.
FT   ZN_FING     457    481       C4-TYPE.
FT   DOMAIN      487    786       STEROID-BINDING.
FT   DOMAIN       48     80       ASP/GLU-RICH (ACIDIC).
FT   VARSPLIC      1    127       MISSING (IN FORM B).
FT   VARSPLIC    452    458       QHNYLCA -> TISYHCS (IN FORM A' AND FORM
FT                                B').
FT   VARSPLIC    459    786       MISSING (IN FORM A' AND FORM B').
FT   CONFLICT     58     58       E -> DD (IN REF. 3).
FT   CONFLICT    480    480       K -> N (IN REF. 3).
FT   CONFLICT    489    489       G -> A (IN REF. 3).
FT   CONFLICT    577    577       R -> T (IN REF. 3).
FT   CONFLICT    642    642       M -> I (IN REF. 3).
SQ   SEQUENCE   786 AA;  85743 MW;  9F84953C CRC32;

  Query Match             65.8%;  Score 352;  DB 7;  Length 786;
  Best Local Similarity   56.1%;  Pred. No. 3.81e-62;
  Matches   37;  Conservative   17;  Mismatches   12;  Indels    0;  Gaps    0;

Db    421 clicqdeasgchygvltcgsckvffkramegqhnylcagrndcivdkirrkncpacrlrk 480
          | :|:| ||| |||| :| :|| ||||:::|::.|.|.. |.| .|| |||.| ||||||
Qy      1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 60

Db    481 ccqagm 486
          | : ||
Qy     61 CYEVGM 66

RESULT   12
```

```
ID   PRGR_MOUSE     STANDARD;      PRT;   923 AA.
AC   Q00175;
DT   01-DEC-1992 (REL. 24, CREATED)
DT   01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   PROGESTERONE RECEPTOR (PR).
GN   PGR OR PR.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 91299759.
RA   SCHOTT D.R., SHYAMALA G., SCHNEIDER W., PARRY G.;
RL   BIOCHEMISTRY 30:7014-7020(1991).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; M68915; G200472; -.
DR   PIR; A39596; A39596.
DR   HSSP; P06536; 1GDC.
DR   PROSITE; PS00031; STEROID_FINGER.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
FT   DOMAIN        1    556       MODULATING, PRO-RICH.
FT   DNA_BIND    557    622       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     557    577       C4-TYPE.
FT   ZN_FING     593    617       C4-TYPE.
FT   DOMAIN      671    923       STEROID-BINDING.
FT   DOMAIN      184    188       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
SQ   SEQUENCE   923 AA;  99073 MW;  EEDFF41A CRC32;

  Query Match             65.8%;  Score 352;  DB 7;  Length 923;
  Best Local Similarity   56.1%;  Pred. No. 3.81e-62;
  Matches   37;  Conservative   17;  Mismatches   12;  Indels    0;  Gaps    0;

Db    557 clicqdeasgchygvltcgsckvffkramegqhnylcagrndcivdkirrkncpacrlrk 616
          | :|:| ||| |||| :| :|| ||||:::|:::|:|:: |:| :|| |||:| ||||||
Qy      1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 60

Db    617 ccqagm 622
          | : ||
Qy     61 CYEVGM 66


RESULT   13
ID   PRGR_RABIT     STANDARD;      PRT;   930 AA.
AC   P06186;
DT   01-JAN-1988 (REL. 06, CREATED)
DT   01-JAN-1988 (REL. 06, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   PROGESTERONE RECEPTOR (PR).
GN   PGR.
OS   ORYCTOLAGUS CUNICULUS (RABBIT).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; LAGOMORPHA.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 87067449.
RA   LOOSFELT H., ATGER M., MISRAHI M., GUIOCHON-MANTEL A., MERIEL C.,
RA   LOGEAT F., BENAROUS R., MILGROM E.;
RL   PROC. NATL. ACAD. SCI. U.S.A. 83:9045-9049(1986).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; M14547; G165632; -.
DR   PIR; A25923; A25923.
DR   HSSP; P06536; 1GDC.
DR   PROSITE; PS00031; STEROID_FINGER.
DR   TRANSFAC; T00697; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
FT   DOMAIN        1    565       MODULATING, PRO-RICH.
FT   DNA_BIND    568    633       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     568    588       C4-TYPE.
FT   ZN_FING     604    628       C4-TYPE.
FT   DOMAIN      678    930       STEROID-BINDING.
FT   DOMAIN      184    188       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
SQ   SEQUENCE   930 AA;  98666 MW;  EE133E86 CRC32;

  Query Match            65.8%;  Score 352;  DB 7;  Length 930;
  Best Local Similarity  56.1%;  Pred. No. 3.81e-62;
  Matches   37;  Conservative   17;  Mismatches  12;  Indels   0;  Gaps   0;

Db    569 clicgdeasgchygvltcgsckvffkramegqhnylcagrndcivdkirrkncpacrlrk 627
          | :||: ||| |||| :| :|| ||||:::|:::|:|:: |:| :|| |||:| ||||||
Qy      1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCIIDKNRRKSCQACRLRK 60

Db    628 ccqagm 633
          | : ||
Qy     61 CYEVGM 66


RESULT   14
ID   PRGR_HUMAN     STANDARD;      PRT;   933 AA.
AC   P06401;
DT   01-JAN-1988 (REL. 06, CREATED)
DT   01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   PROGESTERONE RECEPTOR (PR).
GN   PGR.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 90228361.
RA   KASTNER P., KRUST A., TURCOTTE B., STROPP U., TORA L., GRONEMEYER H.,
RA   CHAMBON P.;
RL   EMBO J. 9:1603-1614(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 87184565.
RA   MISRAHI M., ATGER M., D'AURIOL L., LOOSFELT H., MERIEL C.,
RA   FRIDLANSKY F., GUIOCHON-MANTEL A., GALIBERT F., MILGROM E.;
RL   BIOCHEM. BIOPHYS. RES. COMMUN. 143:740-748(1987).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- ALTERNATIVE PRODUCTS: THERE ARE TWO FORMS OF THE PROGESTERONE
CC       RECEPTOR: FORM A (THAT STARTS WITH MET-1) AND FORM B (THAT STARTS
CC       WITH MET-165).
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; X51730; G35652; -.
DR   EMBL; M15716; G189935; -.
DR   PIR; A03245; QRHUP.
DR   PIR; S09971; S09971.
DR   HSSP; P06536; 1GDC.
DR   MIM; 264080; -.
DR   PROSITE; PS00031; STEROID_FINGER.
DR   TRANSFAC; T00696; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING; PHOSPHORYLATION; ALTERNATIVE SPLICING.
FT   DOMAIN        1    566       MODULATING, PRO-RICH.
FT   DNA_BIND    567    632       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     567    587       C4-TYPE.
FT   ZN_FING     603    627       C4-TYPE.
FT   DOMAIN      681    933       STEROID-BINDING.
FT   DOMAIN      183    187       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT   MOD_RES      41     41       PHOSPHORYLATION (BY CK).
FT   MOD_RES     227    227       PHOSPHORYLATION (BY CK).
FT   MOD_RES     232    232       PHOSPHORYLATION (BY CK).
FT   MOD_RES     552    552       PHOSPHORYLATION (BY CK).
FT   MOD_RES     793    793       PHOSPHORYLATION (BY CK).
FT   VARSPLIC      1    164       MISSING (IN FORM B).
FT   CONFLICT    226    226       S -> G (IN REF. 2).
FT   CONFLICT    256    256       C -> V (IN REF. 2).
FT   CONFLICT    344    344       I -> S (IN REF. 2).
FT   CONFLICT    660    660       L -> V (IN REF. 2).
SQ   SEQUENCE   933 AA;  99041 MW;  CD1124EB CRC32;

  Query Match            65.8%;  Score 352;  DB 7;  Length 933;
  Best Local Similarity  56.1%;  Pred. No. 3.81e-62;
  Matches   37;  Conservative   17;  Mismatches  12;  Indels   0;  Gaps   0;

Db    567 clicgdeasgchygvltcgsckvffkramegqhnylcagrndcivdkirrkncpacrlrk 626
          | :|:| ||| |||| :| :|| ||||:::|:::|:|:: |:| :|| |||:| ||||||
Qy      1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCIIDKNRRKSCQACRLRK 60

Db    627 ccqagm 632
          | : ||
Qy     61 CYEVGM 66


RESULT   15
ID   ANDR_RABIT     STANDARD;      PRT;   709 AA.
AC   P49699;
DT   01-FEB-1996 (REL. 33, CREATED)
DT   01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT   01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE   ANDROGEN RECEPTOR (FRAGMENT).
GN   AR.
OS   ORYCTOLAGUS CUNICULUS (RABBIT).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; LAGOMORPHA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NEW ZEALAND WHITE; TISSUE=PROSTATE;
RX   MEDLINE; 96044663.
RA   KRONGRAD A., WILSON J.D., MCPHAUL M.J.;
RL   J. ANDROL. 16:209-212(1995).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- IN THE ABSENCE OF LIGAND, STEROID HORMONE RECEPTORS ARE THOUGHT
CC       TO BE WEAKLY ASSOCIATED WITH NUCLEAR COMPONENTS; HORMONE BINDING
CC       GREATLY INCREASES RECEPTOR AFFINITY. THE HORMONE-RECEPTOR
CC       COMPLEX APPEARS TO RECOGNIZE DISCRETE DNA SEQUENCES UPSTREAM OF
CC       TRANSCRIPTIONAL START SITES.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; U15366; G577829; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
FT   NON_TER       1      1
FT   DOMAIN       <1    347       MODULATING.
FT   DNA_BIND    349    414       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     349    369       C4-TYPE.
FT   ZN_FING     385    409       C4-TYPE.
FT   DOMAIN      480    709       LIGAND-BINDING.
FT   DOMAIN        3      6       POLY-GLN.
FT   DOMAIN      182    187       POLY-PRO.
FT   DOMAIN      201    207       POLY-ALA.
FT   DOMAIN      254    262       POLY-GLY.
SQ   SEQUENCE   709 AA;  77391 MW;  D8BB7CF3 CRC32;

  Query Match             65.6%;  Score 351;  DB 1;  Length 709;
  Best Local Similarity  59.1%;  Pred. No. 6.71e-62;
  Matches   39;  Conservative   12;  Mismatches  15;  Indels   0;  Gaps   0;

Db     349 clicqdeasgchygaltcgsckvffkraaegkqkylcasrndctidkfrrkncpscrlrk 408
           | :|:| ||| |||   :| :|| ||||: :|    |:|:: |:||||| |||:| :|||||
Qy       1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCIIDKNRRKSCQACRLRK 60

Db     409 cyeagm 414
           ||| ||
Qy      61 CYEVGM 66

Search completed: Thu Dec 18 08:29:28 1997
Job time : 19 secs.
```

```
**************************************************************************
   _   _    ____    ____   ____   ____   _   _
  | \_/ |  |  _ \  / ___| |  _ \ / ___| | | | |
  |     |  | |_) | \___ \ | |_) | |     | |_| |
  | |\/| |  |  __/   ___) ||  _ <| |___  |  _  |
  |_|  |_|  |_|     |____/ |_| \_\\____| |_| |_|    (TM)

**************************************************************************

        Release 2.1D John F. Collins, Biocomputing Research Unit.
        Copyright (c) 1993, 1994, 1995  University of Edinburgh, U.K.
               Distribution rights by IntelliGenetics, Inc

MPsrch_pp   protein - protein database search, using Smith-Waterman algorithm

Run on:          Thu Dec 18 08:39:20 1997;  MasPar time 4.65 Seconds
                                            410.176 Million cell updates/sec
Tabular output not generated

Title:           >US-08-826-361A-3
Description:     (1-66) from US08826361A.pep
Perfect Score:   535
Sequence         1 CAVCSDYASGYHYGVWSCEG            KNRRKSCQACRLRKCYEVGM 66

Scoring table:   PAM 150
                 Gap 11

Searched:        91006 seqs, 28888922 residues

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database:        pir51
                   1:ann1 2:ann2 3:ann3 4:ann4 5:unann1 6:unann2 7:unann3
                   8:unann4 9:unann5 10:unann6 11:unann7 12:unann8
                   13:unann9 14:unann10 15:unenc 16:unrev

Statistics.      Mean 34.170;  Variance 66.867;  scale 0.511

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES
                %
Result        Query
  No   Score  Match Length DB  ID            Description               Pred. No.
-------------------------------------------------------------------------------
    1    530   99.1    586  3  QRXLE         estrogen receptor -       3.34e-87
    2    527   98.5    589  3  QRCHE         estrogen receptor -       1.41e-86
    3    527   98.5    595  3  QRHUE         estrogen receptor -       1.41e-86
    4    527   98.5    599  3  QRMSE         estrogen receptor -       1.41e-86
    5    527   98.5    600  3  QRRTE         estrogen receptor -       1.41e-86
    6    527   98.5    701 16  S64737        90K estrogen recepto      1.41e-86
    7    517   96.6    574  7  A37197        estrogen receptor -       1.70e-84
    8    507   94.8    535  7  S58224        oestrogen receptor -      2.05e-82
    9    387   72.3    433 14  S58087        estrogen receptor re      1.05e-57
   10    387   72.3    433 13  P29345        steroid hormone rece      1.05e-57
   11    374   69.9    521 13  A29345        steroid hormone rece      4.61e-55
   12    352   65.8    786  7  A35466        progesterone recepto      1.31e-50
   13    352   65.8    923 14  I53280        progesterone recepto      1.31e-50
   14    352   65.8    923  7  A39596        progesterone recepto      1.31e-50
   15    352   65.8    930  7  A25923        progesterone recepto      1.31e-50
   16    352   65.8    933  3  QRHUP         progesterone recepto      1.31e-50
   17    351   65.6    899 14  A35995        androgen receptor -       2.08e-50
   18    351   65.6    902 14  B40494        androgen receptor -       2.08e-50
   19    351   65.6    910 13  A34721        androgen receptor A       2.08e-50
   20    351   65.6    911 13  B34721        androgen receptor B       2.08e-50
   21    351   65.6    919 13  A39248        androgen receptor -       2.08e-50
   22    349   65.2    742  3  QRHUGB        glucocorticoid recep      5.26e-50
   23    349   65.2    771  7  A54272        glucocorticoid recep      5.26e-50
   24    349   65.2    776  3  S44047        glucocorticoid recep      5.26e-50
   25    349   65.2    777  3  QRHUGA        glucocorticoid recep      5.26e-50
   26    349   65.2    795  3  QRRTG         glucocorticoid recep      5.26e-50
   27    343   64.1    981 14  A41401        mineralocorticoid re      8.49e-49
   28    343   64.1    984 13  A29513        mineralocorticoid re      8.49e-49
   29    339   63.4    239  7  B23733        progesterone recepto      5.41e-48
   30    339   63.4    783  3  A25691        glucocorticoid recep      5.41e-48
   31    338   63.2    447  7  B34714        retinoic acid recept      8.60e-48
   32    338   63.2    454  7  A33903        retinoic acid recept      8.60e-48
   33    338   63.2    458  7  A34714        retinoic acid recept      8.60e-48
   34    336   62.8    454  7  S06124        retinoic acid recept      2.17e-47
   35    336   62.8    598  7  A37251        probable nuclear hor      2.17e-47
   36    336   62.8    598  7  A61451        steroid hormone rece      2.17e-47
   37    333   62.2    442  7  A38591        retinoic acid recept      8.67e-47
   38    333   62.2    448  7  S02827        retinoic acid recept      8.67e-47
   39    333   62.2    448  7  B56558        retinoic acid recept      8.67e-47
   40    333   62.2    448  7  S05051        retinoic acid recept      8.67e-47
   41    333   62.2    453 13  I50674        retinoic acid recepto     8.67e-47
   42    333   62.2    462  7  A29491        retinoic acid recept      8.67e-47
   43    333   62.2    476  7  B41977        retinoic acid recept      8.67e-47
   44    333   62.2    499 16  I51257        retinoic acid recept      8.67e-47
   45    333   62.2    955 15  S40045        probable transcripti      8.67e-47

                              ALIGNMENTS

RESULT    1
ENTRY              QRXLE      #type complete
TITLE              estrogen receptor - African clawed frog
ORGANISM           #formal_name Xenopus laevis #common_name African clawed frog
DATE               28-Feb-1992 #sequence_revision 14-Jul-1994 #text_change
                     15-Jun-1996
ACCESSIONS         A40907
REFERENCE          A40907
   #authors        Weiler, I.J.; Lew, D.; Shapiro, D.J.
   #journal        Mol. Endocrinol. (1987) 1:355-362
   #title          The Xenopus laevis estrogen receptor: sequence homology with
                     human and avian receptors and identification of multiple
                     estrogen receptor messenger ribonucleic acids.
   #cross-references MUID:90331927
   #accession      A40907
      ##molecule_type mRNA
      ##residues       1-586 ##label WEI
      ##cross-references GB:L20735
COMMENT    The steroid hormones and their receptors are involved in the
             regulation of eukaryotic gene expression and affect cellular
             proliferation and differentiation in target tissues.
COMMENT    In the absence of ligand, steroid hormone receptors are thought to
             be weakly associated with nuclear components; hormone binding
             greatly increases receptor affinity. The hormone-receptor complex
             appears to recognize discrete DNA sequences upstream of
             transcriptional start sites.
CLASSIFICATION   #superfamily estrogen receptor, erbA transforming protein
                   homology
KEYWORDS         DNA binding; nucleus; phosphoprotein; steroid binding;
                   steroid hormone receptor; transcription regulation; zinc
                   finger
FEATURE
   1-174                 #domain amino-terminal #label NH2\
   175-266               #domain DNA binding #status predicted #label DNA\
   179-449               #domain erbA transforming protein homology #label ERBA\
   178-201               #region zinc finger CCCC motif\
   214-237               #region zinc finger CCCC motif\
   251-266               #region nuclear location signal\
   292-544               #domain steroid binding #status predicted #label STE\
   180,183,197,200       #binding_site zinc (Cys) #status predicted\
   216,222,232,235       #binding_site zinc (Cys) #status predicted\
   231                   #binding_site phosphate (Ser) (covalent) #status
                           predicted
SUMMARY          #length 586  #molecular-weight 66080  #checksum 9991
```

```
  Query Match            99.1%;  Score 530;  DB 3;  Length 586;
  Best Local Similarity  98.5%;  Pred. No. 3.34e-97;
  Matches   65;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Db      180 cavcsdyasgyhygvwscegckaffkrsiqghndymcpatnqctidknrrkscqacrlrk 239
            |||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Qy        1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 60

Db      240 cyevgm 245
            ||||||
Qy       61 CYEVGM 66


RESULT     2
ENTRY                 QRCHE       #type complete
TITLE                 estrogen receptor - chicken
ORGANISM              #formal_name Gallus gallus #common_name chicken
DATE                  06-Mar-1992 #sequence_revision 14-Jul-1994 #text_change
                        15-Jun-1996
ACCESSIONS            A40914; S07192
REFERENCE             A40914
   #authors           Maxwell, B.L.; McDonnell, D.P.; Conneely, O.M.; Schulz, T.Z.;
                        Greene, G.L.; O'Malley, B.W.
   #journal           Mol. Endocrinol. (1987) 1:25-35
   #title             Structural organization and regulation of the chicken
                        estrogen receptor
   #cross-references MUID:88318621
   #accession         A40914
      ##status              not compared with conceptual translation
      ##molecule_type mRNA
      ##residues            1-589 ##label MAX
REFERENCE             S07192
   #authors           Krust, A.; Green, S.; Argos, P.; Kumar, V.; Walter, P.;
                        Bornert, J.M.; Chambon, P.
   #journal           EMBO J. (1986) 5:891-897
   #title             The chicken oestrogen receptor sequence: homology with v-erbA
                        and the human oestrogen and glucocorticoid receptors.
   #cross-references MUID:86247578
   #accession         S07192
      ##status              preliminary
      ##molecule_type mRNA
      ##residues            1-256,'F',258-589 ##label KRU
      ##cross-references EMBL:X03805
COMMENT     The steroid hormones and their receptors are involved in the
              regulation of eukaryotic gene expression and affect cellular
              proliferation and differentiation in target tissues.
COMMENT     In the absence of ligand, steroid hormone receptors are thought to
              be weakly associated with nuclear components; hormone binding
              greatly increases receptor affinity. The hormone-receptor complex
              appears to recognize discrete DNA sequences upstream of
              transcriptional start sites.
CLASSIFICATION    #superfamily estrogen receptor; erbA transforming protein
                    homology
KEYWORDS              DNA binding; nucleus; phosphoprotein; steroid binding;
                        steroid hormone receptor; transcription regulation; zinc
                        finger
FEATURE
   1-173                     #domain amino-terminal #label NH2\
   174-265                   #domain DNA binding #status predicted #label DNA\
   177-450                   #domain erbA transforming protein homology #label ERBA\
   177-200                   #region zinc finger CCCC motif\
   213-235                   #region zinc finger CCCC motif\
   250-265                   #region nuclear location signal\
   294-546                   #domain steroid binding #status predicted #label STB\
   179,182,196,199           #binding_site zinc (Cys) #status predicted\
   215,221,231,234           #binding_site zinc (Cys) #status predicted\
   230,299                   #binding_site phosphate (Ser) (covalent) #status
                               predicted
SUMMARY               #length 589  #molecular-weight 66773  #checksum 793

  Query Match            98.5%;  Score 527;  DB 3;  Length 589;
  Best Local Similarity  97.0%;  Pred. No. 1.41e-86;
  Matches   64;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Db      179 cavcndyasgyhygvwscegckaffkrsiqghndymcpatnqctidknrrkscqacrlrk 238
            ||||:||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Qy        1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 60

Db      239 cyevgm 244
            ||||||
Qy       61 CYEVGM 66


RESULT     3
ENTRY                 QRHUE       #type complete
TITLE                 estrogen receptor - human
ALTERNATE_NAMES   hER
ORGANISM              #formal_name Homo sapiens #common_name man
DATE                  28-May-1986 #sequence_revision 28-May-1986 #text_change
                        15-Nov-1996
ACCESSIONS            A94284; A93376; A43021; S27143; S34000; A41925; B41925;
                        A03244; C41925; D41925
REFERENCE             A94284
   #authors           Greene, G.L.; Gilna, P.; Waterfield, M.; Baker, A.; Hort, Y.;
                        Shine, J.
   #journal           Science (1986) 231:1150-1154
   #title             Sequence and expression of human estrogen receptor
                        complementary DNA.
   #cross-references MUID:86122927
   #accession         A94284
      ##molecule_type mRNA
      ##residues            1-595 ##label GR1
      ##cross-references GB:M12674
REFERENCE             A93376
   #authors           Green, S.; Walter, P.; Kumar, V.; Krust, A.; Bornert, J.M.;
                        Argos, P.; Chambon, P.
   #journal           Nature (1986) 320:134-139
   #title             Human oestrogen receptor cDNA: sequence, expression and
                        homology to v-erb-A.
   #cross-references MUID:86146892
   #accession         A93376
      ##molecule_type mRNA
      ##residues            1-595 ##label GR2
      ##cross-references GB:X03635
REFERENCE             A43021
   #authors           Ponglikitmongkol, M.; Green, S.; Chambon, P.
   #journal           EMBO J. (1988) 7:3385-3388
   #title             Genomic organization of the human oestrogen receptor gene.
   #accession         A43021
      ##molecule_type DNA
      ##residues            143-161,215-225,244-254,356-374,'G',376,402-422,447-460,
                              'F',461-467,508-528 ##label PON
REFERENCE             S27140
   #authors           Keaveney, M.; Klug, J.; Gannon, F.
   #journal           DNA Seq. (1992) 2:347-358
   #title             Sequence analysis of the 5' flanking region of the human
                        estrogen receptor gene.
   #accession         S27143
      ##status              preliminary; translation not shown
      ##molecule_type DNA
      ##residues            1-115 ##label KEA
      ##cross-references EMBL:X62462
REFERENCE             S34000
   #authors           Pfeffer, U.; Fecarotta, E.; Castagnetta, L.; Vidali, G.
   #journal           Cancer Res. (1993) 53:741-743
   #title             Estrogen receptor variant messenger RNA lacking exon 4 in
                        estrogen-responsive human breast cancer cell lines.
   #accession         S34000
      ##status              preliminary
      ##molecule_type mRNA
      ##residues            216-254,367-399,'G',401-434 ##label PFE
      ##cross-references EMBL:X73067
      ##note                the authors translated the codon GGG for residue 400 as
                              Val
REFERENCE             A41925
   #authors           Dotzlaw, H.; Alkhalaf, M.; Murphy, L.C.
   #journal      Mol. Endocrinol. (1992) 6:773-785
   #title        Characterization of estrogen receptor variant mRNAs from
                   human breast cancers.
   #cross-references MUID:92293154
   #accession    A41925
      ##molecule_type mRNA
      ##residues       1-214,'ELPTLC' ##label DOT
      ##cross-references NCBIN:106580
      ##experimental_source clone 4; breast cancer
      ##note           sequence has been revised after extraction from NCBI
                         backbone
      ##note           the complete sequence of neither the nucleotide nor the
                         protein is shown in this paper
   #accession    B41925
      ##status         significant sequence differences
      ##molecule_type mRNA
      ##cross-references NCBIN:106597; GB:M69296
      ##experimental_source clone 24; breast cancer
COMMENT    The steroid hormones and their receptors are involved in the
             regulation of eukaryotic gene expression and affect cellular
             proliferation and differentiation in target tissues.
COMMENT    In the absence of ligand, steroid hormone receptors are thought to
             be weakly associated with nuclear components; hormone binding
             greatly increases receptor affinity. The hormone-receptor complex
             appears to recognize discrete DNA sequences upstream of
             transcriptional start sites
GENETICS
   #gene         GDB:ESR
      ##cross-references GDB:119120
   #map_position 6q25.1-6q25.1
   #introns      [illegible]
CLASSIFICATION   #superfamily estrogen receptor; erbA transforming protein
                   homology
KEYWORDS         DNA binding; nucleus; phosphoprotein; steroid hormone
                   receptor; transcription regulation; zinc finger
FEATURE
   1-120              #domain amino-terminal #label NH2\
   121-299            #domain DNA binding #status predicted #label DNA\
   183-456            #domain erbA transforming protein homology #label EPBA\
   185-205            #region zinc finger CCCC motif\
   221-245            #region zinc finger CCCC motif\
   300-595            #domain steroid binding #status predicted #label STB\
   236,305            #binding_site phosphate (Ser) (covalent) #status
                        predicted
SUMMARY          #length 595  #molecular-weight 66258  #checksum 4317

  Query Match             98.5%;  Score 527;  DB 3;  Length 595;
  Best Local Similarity   97.0%;  Pred. No. 1.41e-86;
  Matches   64;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Db     185 cavcndyasgyhygvwscegckaffkrsiqghndymcpatnqctidknrrksgqacrlrk 244
           ||||:||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Qy       1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRLRK 60

Db     245 cyevgm 250
           ||||||
Qy      61 CYEVGM 66


RESULT    4
ENTRY            QRMSE      #type complete
TITLE            estrogen receptor - mouse
ORGANISM         #formal_name Mus musculus #common_name house mouse
DATE             06-Mar-1992 #sequence_revision 14-Jul-1994 #text_change
                   15-Jun-1996
ACCESSIONS       A40061
REFERENCE        A40061
   #authors      White, R.; Lees, J.A.; Needham, M.; Ham, J.; Parker, M.
   #journal      Mol. Endocrinol. (1987) 1:735-744
   #title        Structural organization and expression of the mouse estrogen
                   receptor.
   #cross-references MUID:91042558
   #accession    A40061
      ##molecule_type mRNA
      ##residues       1-599 ##label WHI
      ##cross-references GB:M38651
COMMENT    The steroid hormones and their receptors are involved in the
             regulation of eukaryotic gene expression and affect cellular
             proliferation and differentiation in target tissues.
COMMENT    In the absence of ligand, steroid hormone receptors are thought to
             be weakly associated with nuclear components; hormone binding
             greatly increases receptor affinity. The hormone-receptor complex
             appears to recognize discrete DNA sequences upstream of
             transcriptional start sites.
CLASSIFICATION   #superfamily estrogen receptor; erbA transforming protein
                   homology
KEYWORDS         DNA binding; nucleus; phosphoprotein; steroid binding;
                   steroid hormone receptor; transcription regulation; zinc
                   finger
FEATURE
   1-183              #domain amino terminal #label NH2\
   184-275            #domain DNA binding #status predicted #label DNA\
   187-460            #domain erbA transforming protein homology #label ERBA\
   187-210            #region zinc finger CCCC motif\
   223-245            #region zinc finger CCCC motif\
   260-275            #region nuclear location signal\
   304-556            #domain steroid binding #status predicted #label STB\
   [illegible]        #binding_site zinc (Cys) #status predicted\
   225,231,241,244    #binding_site zinc (Cys) #status predicted\
   240,309            #binding_site phosphate (Ser) (covalent) #status
                        predicted
SUMMARY          #length 599  #molecular-weight [illegible]  #checksum 4749

  Query Match             98.5%;  Score 527;  DB 3;  Length 599;
  Best Local Similarity   97.0%;  Pred. No. 1.41e-86;
  Matches   64;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Db     189 cavcndyasgyhygvwscegckaffkrsiqghndymcpatnqctidknrrksgqacrlrk 248
           ||||:||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Qy       1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRLRK 60

Db     249 cyevgm 254
           ||||||
Qy      61 CYEVGM 66


RESULT    5
ENTRY            QRRTE      #type complete
TITLE            estrogen receptor - rat
ORGANISM         #formal_name Rattus norvegicus #common_name Norway rat
DATE             30-Sep-1991 #sequence_revision 14-Jul-1994 #text_change
                   15-Jun-1996
ACCESSIONS       S07379; S16731
REFERENCE        S07379
   #authors      Koike, S.; Sakai, M.; Muramatsu, M.
   #journal      Nucleic Acids Res. (1987) 15:2499-2513
   #title        Molecular cloning and characterization of rat estrogen
                   receptor cDNA.
   #cross-references MUID:87174780
   #accession    S07379
      ##molecule_type mRNA
      ##residues       1-600 ##label KOI
      ##cross-references EMBL:Y00102
REFERENCE        S16731
   #authors      Maggi, A.M.A.
   #submission   submitted to the EMBL Data Library, June 1991
   #accession    S16731
      ##molecule_type mRNA
      ##residues       1-487,'T',489-600 ##label MAG
      ##cross-references EMBL:X61098
COMMENT    The steroid hormones and their receptors are involved in the
             regulation of eukaryotic gene expression and affect cellular
             proliferation and differentiation in target tissues.
COMMENT    In the absence of ligand, steroid hormone receptors are thought to
             be weakly associated with nuclear components; hormone binding
             greatly increases receptor affinity. The hormone-receptor complex
```

```
               appears to recognize discrete DNA sequences upstream of
               transcriptional start sites.
CLASSIFICATION   #superfamily estrogen receptor, erbA transforming protein
                   homology
KEYWORDS         DNA binding; nucleus; phosphoprotein; steroid binding;
                   steroid hormone receptor; transcription regulation; zinc
                   finger
FEATURE
   1-184               #domain amino-terminal #label NH2\
   185-276             #domain DNA binding #status predicted #label DNA\
   188-461             #domain erbA transforming protein homology #label ERBA\
   188-211             #region zinc finger CCCC motif\
   224-246             #region zinc finger CCCC motif\
   261-276             #region nuclear location signal\
   305-557             #domain steroid binding #status predicted #label STR\
   190,193,207,210     #binding_site zinc (Cys) #status predicted\
   226,232,242,245     #binding_site zinc (Cys) #status predicted\
   241,310             #binding_site phosphate (Ser) (covalent) #status
                         predicted
SUMMARY          #length 600  #molecular-weight 67029  #checksum 4599

  Query Match             98.5%;  Score 527;  DB 3;  Length 600;
  Best Local Similarity   97.0%;  Pred. No. 1.41e-86;
  Matches    64;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Db     190 cavcndyasgyhygvwscegckaffkrsiqghndymcpatnqctidknrrkscqacrlrk 249
           ||||:|||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Qy       1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 60

Db     250 cyevgm 255
           ||||||
Qy      61 CYEVGM 66


RESULT     6
ENTRY            S64737     #type complete
TITLE            80K estrogen receptor - human
ORGANISM         #formal_name Homo sapiens #common_name man
DATE             28 Oct 1996 #sequence_revision 28 Oct 1996 #text_change
                   28-Oct-1996
ACCESSIONS       S64737
REFERENCE        S64737
   #authors      Pink, J.J.; Wu, S.Q.; Wolf, D.M.; Bilimoria, M.M.; Jordan,
                   V.C.
   #journal      Nucleic Acids Res. (1996) 24:962-969
   #title        A novel 80 kDa human estrogen receptor containing a
                   duplication of exons 6 and 7.
   #accession    S64737
      ##status          preliminary
      ##residues        1-701 ##label PIN
      ##cross-references EMBL:U47678
SUMMARY          #length 701  #molecular-weight 78430  #checksum 4951

  Query Match             98.5%;  Score 527;  DB 16;  Length 701;
  Best Local Similarity   97.0%;  Pred. No. 1.41e-86;
  Matches    64;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Db     185 cavcndyasgyhygvwscegckaffkrsiqghndymcpatnqctidknrrkscqacrlrk 244
           ||||:|||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Qy       1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 60

Db     245 cyevgm 250
           ||||||
Qy      61 CYEVGM 66


RESULT     7
ENTRY            A37197     #type complete
TITLE            estrogen receptor - rainbow trout
ORGANISM         #formal_name Oncorhynchus mykiss #common_name rainbow trout
DATE             31-Jan-1992 #sequence_revision 31 Jan 1992 #text_change
                   03-Jun-1996
ACCESSIONS       A37197; A40070
REFERENCE        A37197
   #authors      Pakdel, F.; Le Gac, F.; Le Goff, P.; Valotaire, Y.
   #journal      Mol. Cell. Endocrinol. (1990) 71:195-204
   #title        Full-length sequence and in vitro expression of rainbow trout
                   estrogen receptor cDNA.
   #cross-references MUID:91006824
   #accession    A37197
      ##status          preliminary
      ##molecule_type mRNA
      ##residues        1-574 ##label PAK
      ##cross-references GB:M31559
REFERENCE        A40070
   #authors      Pakdel, F.; Le Guellec, C.; Vaillant, C.; Le Roux, M.G.;
                   Valotaire, Y.
   #journal      Mol. Endocrinol. (1989) 3:44-51
   #title        Identification and estrogen induction of two estrogen
                   receptors (ER) messenger ribonucleic acids in the rainbow
                   trout liver: sequence homology with other ERs.
   #cross-references MUID:89127284
   #accession    A40070
      ##status          preliminary
      ##molecule_type mRNA
      ##residues        150-574 ##label PA2
      ##cross-references GB:M31559
CLASSIFICATION   #superfamily estrogen receptor; erbA transforming protein
                   homology
KEYWORDS         DNA binding; nucleus; steroid hormone receptor; zinc finger
FEATURE
   145-421             #domain erbA transforming protein homology #label ERBA\
   147-167             #region zinc finger\
   183-207             #region zinc finger
SUMMARY          #length 574  #molecular weight 62871  #checksum 1057

  Query Match             96.6%;  Score 517;  DB 7;  Length 574;
  Best Local Similarity   93.9%;  Pred. No. 1.70e-84;
  Matches    62;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Db     147 cavcsdfasgyhygvwscegckaffkrsiqghndymcpatnqctmdrnrrkscqacrlrk 206
           ||||||:|||||||||||||||||||||||||||||:|||||||:|:|||||||||||||
Qy       1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 60

Db     207 cyevgm 212
           ||||||
Qy      61 CYEVGM 66


RESULT     8
ENTRY            S58224     #type fragment
TITLE            oestrogen receptor - Atlantic salmon (fragment)
ORGANISM         #formal_name Salmo salar #common_name Atlantic salmon
DATE             13 Jan 1996 #sequence_revision 01-Mar 1996 #text_change
                   26-Apr-1996
ACCESSIONS       S58224
REFERENCE        S58224
   #authors      Rogers, S.A.; Llewellyn, L.; Ramsurn, V.P.; Sweeney, G.E.;
                   Wigham, T.
   #submission   submitted to the EMBL Data Library, July 1995
   #accession    S58224
      ##status          preliminary
      ##molecule_type mRNA
      ##residues        1-535 ##label ROG
      ##cross-references EMBL X89959
CLASSIFICATION   #superfamily erbA transforming protein homology
FEATURE
   103-382             #domain erbA transforming protein homology #label ERBA
SUMMARY          #length 535  #checksum 9963

  Query Match             94.8%;  Score 507;  DB 7;  Length 535;
  Best Local Similarity   92.4%;  Pred. No. 2.05e-82;
  Matches    61;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Db     105 cavcsdfasgyhygfwscegckaffkrsiqghndymcpatnqctmdrnrrkscqacrlrk 164
           ||||||:||||||||| ||||||||||||||||||||:|||||||:|:|||||||||||||
Qy       1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRLRK 60

Db     165 cyevgm 170
           ||||||
Qy      61 CYEVGM 66


RESULT    9
ENTRY             S58087     #type complete
TITLE             estrogen receptor related protein - mouse
ORGANISM          #formal_name Mus musculus #common_name house mouse
DATE              13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change
                    26-Apr-1996
ACCESSIONS        S58087
REFERENCE         S58087
   #authors       Pettersson, K.; Svensson, K.; Mattsson, R.; Carlsson, B.;
                    Ohlsson, R.; Berkenstam, A.
   #submission    submitted to the EMBL Data Library, July 1995
   #description   Expression of a novel member of estrogen response
                    element-binding nuclear receptors is restricted to the
                    early stages of chorion formation during mouse
                    embryogenesis.
   #accession     S58087
      ##status          preliminary
      ##molecule_type mRNA
      ##residues        1-433 ##label PET
      ##cross-references EMBL:X89594
CLASSIFICATION    #superfamily erbA transforming protein homology
FEATURE
   101-352              #domain erbA transforming protein homology #label EFBA
SUMMARY           #length 433  #molecular-weight 48400  #checksum 1075

  Query Match             72.3%;  Score 387;  DB 14;  Length 433;
  Best Local Similarity  72.7%;  Pred. No. 1.05e-57;
  Matches    48;  Conservative    7;  Mismatches  11;  Indels    0;  Gaps    0;

Db     103 clvcgdiasgyhygvasceackaffkrtiqgnieyscpatneceitkrrrkscqacrfmk 162
           | |.:| |||||||| |||:||||||||:|||: :| ||||||:| | | |||||||||: |
Qy       1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRLRK 60

Db     163 clkvgm 168
           |  |||
Qy      61 CYEVGM 66


RESULT   10
ENTRY             B29345     #type complete
TITLE             steroid hormone receptor ERR2 precursor - human
ORGANISM          #formal_name Homo sapiens #common_name man
DATE              15-Dec-1988 #sequence_revision 15-Dec-1988 #text_change
                    03-Jun-1996
ACCESSIONS        B29345
REFERENCE         A29345
   #authors       Giguere, V.; Yang, N.; Segui, P.; Evans, R.M.
   #journal       Nature (1988) 331:91-94
   #title         Identification of a new class of steroid hormone receptors.
   #cross-references MUID:88122546
   #accession     B29345
      ##molecule_type mRNA
      ##residues        1-433 ##label GIG
      ##cross-references GB:X51417; GB:Y00290
CLASSIFICATION    #superfamily erbA transforming protein homology
KEYWORDS          DNA binding; zinc finger
FEATURE
   101-352              #domain erbA transforming protein homology #label EFBA\
   103-123              #region zinc finger\
   139-163              #region zinc finger
SUMMARY           #length 433  #molecular-weight 48297  #checksum 1133

  Query Match             72.3%;  Score 387;  DB 13;  Length 433;
  Best Local Similarity  72.7%;  Pred. No. 1.05e-57;
  Matches    48;  Conservative    7;  Mismatches  11;  Indels    0;  Gaps    0;
```

```
Db     103 clvcgdiasgyhygvasceackaffkrtiqgnieyscpatneceitkrrrkscqacrfmk 162
           | ||:| |||||||| |||:||||||||:|||: :| ||||||:| ' | |||||||||: |
Qy       1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRLRK 60

Db     163 clkvgm 168
           |  |||
Qy      61 CYEVGM 66


RESULT   11
ENTRY             A29345     #type complete
TITLE             steroid hormone receptor ERR1 precursor - human
ALTERNATE_NAMES   estrogen-related receptor
ORGANISM          #formal_name Homo sapiens #common_name man
DATE              15-Dec-1988 #sequence_revision 15-Dec-1988 #text_change
                    03-Jun-1996
ACCESSIONS        A29345; A49074
REFERENCE         A29345
   #authors       Giguere, V.; Yang, N.; Segui, P.; Evans, R.M.
   #journal       Nature (1988) 331:91-94
   #title         Identification of a new class of steroid hormone receptors.
   #cross-references MUID:88122546
   #accession     A29345
      ##molecule_type mRNA
      ##residues        1-521 ##label GIG
      ##cross references EMBL:X51416; EMBL:Y00290
REFERENCE         A49074
   #authors       Wiley, S.R.; Kraus, R.J.; Zuo, F.; Murray, E.E.; Loritz, K.;
                    Mertz, J.E.
   #journal       Genes Dev. (1993) 7:2206-2219
   #title         SV40 early-to-late switch involves titration of cellular
                    transcriptional repressors.
   #accession     A49074
      ##status          preliminary
      ##molecule_type protein
      ##residues        166-169,'X',171-173 ##label WIL
CLASSIFICATION    #superfamily erbA transforming protein homology
KEYWORDS          DNA binding; nucleus; zinc finger
FEATURE
   174-434              #domain erbA transforming protein homology #label EFBA\
   176-196              #region zinc finger\
   212-236              #region zinc finger
SUMMARY           #length 521  #molecular-weight 55742  #checksum 5840

  Query Match             69.9%;  Score 374;  DB 13;  Length 521;
  Best Local Similarity  69.7%;  Pred. No. 4.61e-55;
  Matches    46;  Conservative    8,  Mismatches  12,  Indels    0,  Gaps    0;

Db     176 clvcgdvasgyhygvasceackaffkrtiqgsieyscpasneceitkrrrkacqacrftk 235
           | ||.| |||||||| |||:||||||||:|||  :| |||:|:| | | |||:|'|'|: |
Qy       1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRLRK 60

Db     236 clrvgm 241
           |  |||
Qy      61 CYEVGM 66


RESULT   12
ENTRY             A35466     #type complete
TITLE             progesterone receptor form B - chicken
CONTAINS          progesterone receptor form A
ORGANISM          #formal_name Gallus gallus #common_name chicken
DATE              18-Nov-1994 #sequence_revision 18-Nov-1994 #text_change
                    03-Jun-1996
ACCESSIONS        A35466; S06294; A40903; A34661; A34312; A40911
REFERENCE         A35466
   #authors       Jeltsch, J.M.; Turcotte, B.; Garnier, J.M.; Lerouge, T.;
                    Krozowski, Z.; Gronemeyer, H.; Chambon, P.
   #journal       J. Biol. Chem. (1990) 265:3967-3974
   #title         Characterization of multiple mRNAs originating from the
                    chicken progesterone receptor gene. Evidence for a specific
                    transcript encoding form A.
   #accession     A35466
```

```
      ##status          translation not shown
      ##molecule_type DNA
      ##residues        1-786 ##label JEL
      ##cross-references GB:J05240
REFERENCE          S06284
   #authors        Gronemeyer, H.; Turcotte, B.; Quirin Stricker, C.; Bocquel,
                     M.T.; Meyer, M.E.; Krozowski, Z.; Jeltsch, J.M.; Lerouge,
                     T.; Garnier, J.M.; Chambon, P.
   #journal        EMBO J. (1987) 6:3985-3994
   #title          The chicken progesterone receptor: sequence, expression and
                     functional analysis.
   #cross-references MUID:88166640
   #accession      S06284
      ##molecule_type DNA
      ##residues        1-786 ##label GRO
      ##cross-references EMBL:Y00092
REFERENCE          A40903
   #authors        Conneely, O.M.; Dobson, A.D.W.; Tsai, M.J.; Beattie, W.G.;
                     Toft, D.O.; Huckaby, C.S.; Zarucki, T.; Schrader, W.T.;
                     O'Malley, B.W.
   #journal        Mol. Endocrinol. (1987) 1:517-525
   #title          Sequence and expression of a functional chicken progesterone
                     receptor.
   #cross-references MUID:91042592
   #accession      A40903
      ##status          preliminary
      ##molecule_type mRNA
      ##residues        1-64,'E',65-786 ##label CON
      ##cross-references GB:M37518
REFERENCE          A24661
   #authors        Conneely, O.M.; Sullivan, W.P.; Toft, D.O.; Birnbaumer, M.;
                     Cook, R.G.; Maxwell, B.L.; Zarucki-Schulz, T.; Greene,
                     G.L.; Schrader, W.T.; O'Malley, B.W.
   #journal        Science (1986) 233:767-770
   #title          Molecular cloning of the chicken progesterone receptor.
   #cross-references MUID:86289413
   #accession      A24661
      ##molecule_type mRNA
      ##residues        128-133,'E',135-147,'E',149-164 ##label CO2
      ##note            the authors translated the codon CAG for residue 7 as
                          Glu and CAG for residue 21 as Glu
REFERENCE          A24312
   #authors        Jeltsch, J.M.; Krozowski, Z.; Quirin-Stricker, C.;
                     Gronemeyer, H.; Simpson, R.J.; Garnier, J.M.; Krust, A.;
                     Jacob, F.; Chambon, P.
   #journal        Proc. Natl. Acad. Sci. U.S.A. (1986) 83:5424-5428
   #title          Cloning of the chicken progesterone receptor.
   #cross-references MUID:86287271
   #accession      A24312
      ##molecule_type mRNA
      ##residues        417-490 ##label JE2
      ##cross-references GB:M14280
      ##note            amino acid and corresponding nucleotide sequences are
                          also shown for three smaller peptides
REFERENCE          A40911
   #authors        Birnbaumer, M.; Hinrichs-Rosello, M.V.; Cook, R.G.; Schrader,
                     W.T.; O'Malley, B.W.
   #journal        Mol. Endocrinol. (1987) 1:249-259
   #title          Chemical and antigenic properties of pure 108,000 molecular
                     weight chick progesterone receptor.
   #cross-references MUID:88288199
   #accession      A40911
      ##status          preliminary
      ##molecule_type protein
      ##residues        128-133,'E',135-147,'E',149-164,546-558 ##label BIF
GENETICS
   #introns        405/1; 451/1; 4[illegible]/1; 541/1; [illegible]/2; [illegible]/1; 735/3
CLASSIFICATION     #superfamily progesterone receptor; erbA transforming protein
                     homology
KEYWORDS           DNA binding; nucleus; steroid hormone receptor; zinc finger
FEATURE
   1-786                #product progesterone receptor form B #status predicted
                          #label MA1\
   128-786              #product progesterone receptor form A #status predicted
                          #label MA2\
   419-682              #domain erbA transforming protein homology #label ERBA\
   421-441              #region zinc finger\
   457-481              #region zinc finger
SUMMARY            #length 786  #molecular-weight 85743  #checksum 8923

  Query Match             65.8%;  Score 352;  DB 7;  Length 786;
  Best Local Similarity   56.1%;  Pred. No. 1.31e-50;
  Matches   37;  Conservative   17;  Mismatches   12;  Indels    0;  Gaps    0;

Db        421 cliqdeasgchygvltcgsckvffkramegqhnylcagrndcivdkirrkncpacrlrk 480
              | :|:| ||| |||| :| :|| ||||:::|:::|:|:: |:| :|| |||:| ||||||
Qy          1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 60

Db        481 ccqagm 486
              | : ||
Qy         61 CYEVGM 66


RESULT   13
ENTRY              I53280     #type complete
TITLE              progesterone receptor - rat
ORGANISM           #formal_name Rattus norvegicus #common_name Norway rat
DATE               02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change
                     02-Aug-1996
ACCESSIONS         I53280
REFERENCE          I53280
   #authors        Park-Sarge, O.K.; Mayo, K.E.
   #journal        Endocrinology (1994) 134:709-718
   #title          Regulation of the progesterone receptor gene by gonadotropins
                     and cyclic adenosine 3',5'-monophosphate in rat granulosa
                     cells.
   #cross-references MUID:94130817
   #accession      I53280
      ##status          preliminary; translated from GB/EMBL/DDBJ
      ##molecule_type mRNA
      ##residues        1-923 ##label RES
      ##cross-references GB:L16922; NID:g463282; CDS_PID:g463283
SUMMARY            #length 923  #molecular-weight 99407  #checksum 9843

  Query Match             65.8%;  Score 352;  DB 14;  Length 923;
  Best Local Similarity   56.1%;  Pred. No. 1.31e-50;
  Matches   37;  Conservative   17;  Mismatches   12;  Indels    0;  Gaps    0;

Db        557 cliqdeasgchygvltcgsckvffkramegqhnylcagrndcivdkirrkncpacrlrk 616
              | :|:| ||| |||| :| :|| ||||:::|:::|:|:: |:| :|| |||:| ||||||
Qy          1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 60

Db        617 ccqagm 622
              | : ||
Qy         61 CYEVGM 66


RESULT   14
ENTRY              A39596     #type complete
TITLE              progesterone receptor B form - mouse
ORGANISM           #formal_name Mus musculus #common_name house mouse
DATE               20-Mar-1992 #sequence_revision 20-Mar-1992 #text_change
                     03-Jun-1996
ACCESSIONS         A39596
REFERENCE          A39596
   #authors        Schott, D.R.; Shyamala, G.; Schneider, W.; Parry, G.
   #journal        Biochemistry (1991) 30:7014-7020
   #title          Molecular cloning, sequence analyses, and expression of
                     complementary DNA encoding murine progesterone receptor.
   #cross-references MUID:91299759
   #accession      A39596
      ##status          preliminary
      ##molecule_type mRNA
      ##residues        1-923 ##label SCH
      ##cross-references GB:J05333
CLASSIFICATION     #superfamily progesterone receptor; erbA transforming protein
```

```
                 homology
KEYWORDS         DNA binding; nucleus; steroid hormone receptor; zinc finger
FEATURE
   555-819            #domain erbA transforming protein homology #label ERBA\
   557-577            #region zinc finger\
   593-617            #region zinc finger
SUMMARY          #length 923  #molecular-weight 99073  #checksum 6627

  Query Match             65.8%;  Score 352;  DB 7;  Length 923;
  Best Local Similarity  56.1%;  Pred. No. 1.31e-50;
  Matches    37;  Conservative   17;  Mismatches  12;  Indels   0;  Gaps   0;

Db     557 clicqdeasgchygvltcgsckvffkramegqhnylcagrndcivdkirrkncpacrlrk 616
           | :|:| ||| |||| :| :|| |||::::|:::|:|:: |:| :|| |||:| ||||||
Qy       1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 60

Db     617 ccqagm 622
           | : ||
Qy      61 CYEVGM 66


RESULT   15
ENTRY            A25923     #type complete
TITLE            progesterone receptor - rabbit
ORGANISM         #formal_name Oryctolagus cuniculus #common_name domestic
                   rabbit
DATE             16-Aug-1988 #sequence_revision 16-Aug-1988 #text_change
                   03-Jun-1996
ACCESSIONS       A25923
REFERENCE        A25923
   #authors      Loosfelt, H.; Atger, M.; Misrahi, M.; Guiochon-Mantel, A.;
                   Meriel, C.; Logeat, F.; Benarous, R.; Milgrom, E.
   #journal      Proc. Natl. Acad. Sci. U.S.A. (1986) 83:9045-9049
   #title        Cloning and sequence analysis of rabbit progesterone-receptor
                   complementary DNA.
   #cross-references MUID:87067449
   #accession    A25923
      ##molecule_type mRNA
      ##residues      1-930 ##label LOO
CLASSIFICATION   #superfamily progesterone receptor; erbA transforming protein
                   homology
KEYWORDS         DNA binding; nucleus; steroid hormone receptor; zinc finger
FEATURE
   566-826            #domain erbA transforming protein homology #label ERBA\
   568-588            #region zinc finger\
   604-628            #region zinc finger
SUMMARY          #length 930  #molecular-weight 98666  #checksum 2858

  Query Match             65.8%;  Score 352;  DB 7;  Length 930;
  Best Local Similarity  56.1%;  Pred. No. 1.31e-50;
  Matches    37;  Conservative   17;  Mismatches  12;  Indels   0;  Gaps   0;

Db     568 clicqdeasgchygvltcgsckvffkramegqhnylcagrndcivdkirrkncpacrlrk 627
           | :|:| ||| |||| :| :|| |||::::|:::|:|:: |:| :|| |||:| ||||||
Qy       1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 60

Db     628 ccqagm 633
           | : ||
Qy      61 CYEVGM 66

Search completed. Thu Dec 18 08:28:52 1997
Job time : 32 secs.
```

```
*****************************************************************

                         MPSRCH (TM)

*****************************************************************

        Release 2.1D John F. Collins, Biocomputing Research Unit
        Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
               Distribution rights by IntelliGenetics, Inc.

MPsrch_pp   protein - protein database search, using Smith-Waterman algorithm

Run on:          Thu Dec 18 08:23:46 1997; MasPar time 1.71 Seconds
                                           187.124 Million cell updates/sec
Tabular output not generated.

Title:           >US-08-826-361A-3
Description:     (1-66) from US08826361A.pep
Perfect Score:   535
Sequence:        1 CAVCSDYASGYHYGVWSCEG          KNRRKSCQACRLRKCYEVGM 66

Scoring table:   PAM 150
                 Gap 11

Searched:        54101 seqs, 4836506 residues

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database:        a-issued
                   1:back1 2:51 3:52 4:53 5:54 6:55 7:56 8:PCT90 9:PCT91
                   10:PCT92 11:PCT93 12:PCT94 13:PCT95 14:PCT96

Statistics:      Mean 24.490;  Variance 94.838;  scale 0.258

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed
      and is derived by analysis of the total score distribution.

                              SUMMARIES
              %
Result       Query
  No.  Score Match Length DB  ID          Description           Pred. No.
-------------------------------------------------------------------------
     1   527  98.5    410  1  5223606-5   Patent No. 5223606    5.18e-46
     2   352  65.8    393  1  5223606-6   Patent No. 5223606    2.54e-27
     3   349  65.2     66  1  5217867-1   Patent No. 5217867.   5.26e-27
     4   349  65.2     89  6  US-08-561-  Sequence 2, Applicatio 5.26e-27
     5   349  65.2     89  5  US-08-087-  Sequence 2, Applicatio 5.26e-27
     6   349  65.2    356  1  5223606-7   Patent No. 5223606    5.26e-27
     7   349  65.2    795  3  US-07-716-  Sequence 5, Applicatio 5.26e-27
     8   338  63.2    474  1  5260432-2   Patent No. 5260432.   7.03e-26
     9   333  62.2    368  1  5223606-3   Patent No. 5223606.   2.55e-25
    10   333  62.2    448  1  5223606-2   Patent No. 5223606    2.55e-25
    11   333  62.2    462 10  PCT-US92-0  Sequence 4, Applicatio 2.55e-25
    12   333  62.2    797 10  PCT-US92-0  Sequence 2, Applicatio 2.55e-25
    13   320  59.8    440  6  US-08-333-  Sequence 8, Applicatio 5.93e-24
    14   317  59.3    452  1  5171671-2   Patent No. 5171671.   1.22e-23
    15   310  57.9    451 13  PCT-US95-1  Sequence 2, Applicatio 6.63e-23
    16   310  57.9    469 13  PCT-US95-1  Sequence 2, Applicatio 6.63e-23
    17   310  57.9    484 13  PCT-US95-1  Sequence 1, Applicatio 6.63e-23
    18   307  57.4    549  6  US-08-333-  Sequence 2, Applicatio 1.37e-22
    19   303  56.6     70  3  US-07-737-  Sequence 3, Applicatio 3.58e-22
    20   302  56.4    446  5  US-07-952-  Sequence 4, Applicatio 4.55e-22
    21   302  56.4    523  5  US-07-952-  Sequence 2, Applicatio 4.55e-22
    22   299  55.9    452  9  PCT-US91-0  Sequence 2, Applicatio 9.37e-22
    23   299  55.9    467  9  PCT-US91-0  Sequence 4, Applicatio 9.37e-22
    24   295  55.1    462  9  PCT-US91-0  Sequence 6, Applicatio 2.45e-21
    25   291  54.4    443  7  US-08-342-  Sequence 4, Applicatio 6.41e-21
    26   291  54.4    446 13  PCT-US95-1  Sequence 3, Applicatio 6.41e-21
    27   291  54.4    513  7  US-08-464-  Sequence 2, Applicatio 6.41e-21
    28   287  53.6    601  6  US-08-333-  Sequence 14, Applicati 1.67e-20
    29   284  53.1    461 13  PCT-US95-1  Sequence 2, Applicatio 3.43e-20
    30   284  53.1    461 13  PCT-US95-1  Sequence 2, Applicatio 3.43e-20
    31   284  53.1    461  7  US-08-330-  Sequence 2, Applicatio 3.43e-20
    32   284  53.1    461  7  US-08-330-  Sequence 2, Applicatio 3.43e-20
    33   280  52.3    460  7  US-08-342-  Sequence 2, Applicatio 8.95e-20
    34   278  52.0    558  6  US-08-333   Sequence 12, Applicati 1.44e-19
    35   277  51.8    355  1  5223606-4   Patent No. 5223606.   1.83e-19
    36   273  51.0    410  1  5436126-1   Patent No. 5436126.   4.77e-19
    37   265  49.5    439  6  US-08-333-  Sequence 10, Applicati 3.22e-18
    38   249  46.5    433 12  PCT-US94-0  Sequence 2, Applicatio 1.44e-16
    39   230  43.0     89  5  US-08-087-  Sequence 2, Applicatio 1.28e-14
    40   230  43.0     89  6  US-08-561-  Sequence 2, Applicatio 1.28e-14
    41   160  29.9    367  3  US-07-737-  Sequence 4, Applicatio 1.36e-07
    42    81  15.1    303 11  PCT-US93-0  Sequence 1, Applicatio 1.33e+00
    43    74  13.8     59  7  US-08-358-  Sequence 15, Applicati 1.15e+01
    44    72  13.5    474  1  5395760-4   Patent No. 5395760    1.69e+01
    45    71  13.3    335  7  US-08-421-  Sequence 2, Applicatio 2.04e+01


                              ALIGNMENTS

RESULT    1
ID   5223606-5      STANDARD,      PRT,    444 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Patent No. 5223606.
XX
CC   Patent No. 5223606
CC        APPLICANT:  BLADUIN DE THE, HUGHES;MARCHIO, AGNES;TIOLLAIS,
CC   PIERRE;DEJEAN, ANNE
CC        TITLE OF INVENTION: STEROID/THYROID HORMONE RECEPTOR-RELATED
CC   PROTEIN INAPPROPRIATELY EXPRESSED IN HUMAN HEPATOCELLULAR CARCINOMA
CC        NUMBER OF SEQUENCES: 11
CC        CURRENT APPLICATION DATA:
CC          APPLICATION NUMBER: US/07/134,130
CC          FILING DATE: 17-DEC-1987
CC        PRIOR APPLICATION DATA:
CC   SEQ ID NO:5:
CC          LENGTH: 410
SQ   SEQUENCE   444 AA; 50020 MW; 1110634 CN;

  Query Match              98.5%;  Score 527;  DB 1;  Length 410;
  Best Local Similarity  97.0%;  Pred. No. 5.18e-46;
  Matches   64;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Db        1 CAVCNDYASGYHYGVWSCEGCKAFFKRSIQGHNDYMCPATNQCTIDKNRRKSCQACRLRK 60
            ||||:||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Qy        1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 60

Db       61 CYEVGM 66
            ||||||
Qy       61 CYEVGM 66


RESULT    2
ID   5223606-6      STANDARD,      PRT,    393 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Patent No. 5223606.
XX
CC   Patent No. 5223606
CC       APPLICANT:  BLAUDIN DE THE, HUGHES;MARCHIO, AGNES;TIOLLAIS,
CC    PIERRE;DEJEAN, ANNE
CC       TITLE OF INVENTION: STEROID/THYROID HORMONE RECEPTOR-RELATED
CC    PROTEIN INAPPROPRIATELY EXPRESSED IN HUMAN HEPATOCELLULAR CARCINOMA
CC       NUMBER OF SEQUENCES: 11
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/134,130
CC         FILING DATE: 17-DEC-1987
CC       PRIOR APPLICATION DATA:
CC    SEQ ID NO:6:
CC         LENGTH: 363
SQ    SEQUENCE   393 AA; 44638 MW; 907537 CN;

  Query Match             65.8%;  Score 352;  DB 1;  Length 363;
  Best Local Similarity   56.1%;  Pred. No. 2.54e-27;
  Matches   37;  Conservative   17;  Mismatches   12;  Indels    0;  Gaps    0;

Db        1 CLICGDEASGCHYGVLTCGSCKVFFKRAMEGQHNYLCAGRNDCIVDKIRRKNCPACRLRK 60
            | :|:| ||| |||| :| :|| ||||...|...|.|.. |.| .|| |||.| ||||||
Qy        1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRLRK 60

Db       61 CCQAGM 66
            | : ||
Qy       61 CYEVGM 66


RESULT    3
ID   5217867-1      STANDARD;      PRT;    71 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Patent No. 5217867.
XX
CC   Patent No. 5217867
CC       APPLICANT: EVANS, RONALD M.;HOLLENBERG, STANLEY M.
CC       TITLE OF INVENTION: RECEPTORS THEIR IDENTIFICATION,
CC    CHARACTERIZATION, PREPARATION AND USE
CC       NUMBER OF SEQUENCES: 4
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/278,614
CC         FILING DATE: 30-NOV-1988
CC    SEQ ID NO:1:
CC         LENGTH: 66
SQ    SEQUENCE   71 AA; 7915 MW; 22607 CN;

  Query Match             65.2%;  Score 349;  DB 1;  Length 66;
  Best Local Similarity   59.1%;  Pred. No. 5.26e-27;
  Matches   39;  Conservative   14;  Mismatches   13;  Indels    0;  Gaps    0;

Db        1 CLVCSDEASGCHYGVLTCGSCKVFFKRAVEGQHNYLCAGRNDCIIDKIRRKNCPACRYRK 60
            | |||| ||| |||| :| :|| ||||:::|:::|:|:: |:| ||| |||:| ||| ||
Qy        1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRLRK 60

Db       61 CLQAGM 66
            | : ||
Qy       61 CYEVGM 66


RESULT    4
ID   US-08-561-072-3      STANDARD;      PRT;    89 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 3, Application US/08561072
XX
CC    Sequence 3, Application US/08561072
CC    Patent No. 5571791
CC     GENERAL INFORMATION:
CC       APPLICANT:  E. Brad Thompson
```

```
CC       APPLICANT:  Lynne V. Nazareth
CC       TITLE OF INVENTION:  ANTI-TUMOR ACTIVITY OF A MODIFIED FRAGMENT
CC       TITLE OF INVENTION:  OF THE GLUCOCORTICOID RECEPTOR
CC       NUMBER OF SEQUENCES:  3
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Arnold, White & Durkee
CC         STREET:  P.O. Box 4433
CC         CITY:  Houston
CC         STATE:  Texas
CC         COUNTRY:  USA
CC         ZIP:  77210
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  WordPerfect 5.1
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/561,072
CC         FILING DATE:  20-NOV-1995
CC         CLASSIFICATION:  530
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  08/087,151
CC         FILING DATE:  July 7, 1993
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Hodgins, Daniel S.
CC         REGISTRATION NUMBER:  31,026
CC         REFERENCE/DOCKET NUMBER:  UTSG:141/HOD
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  512-320-7200
CC         TELEFAX:  512-474-7577
CC         TELEX:  No. 5571791 Applicable
CC     INFORMATION FOR SEQ ID NO:  3:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  89 amino acid residues
CC         TYPE:  amino acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  peptide
SQ    SEQUENCE   89 AA; 9644 MW; 34324 CN;

  Query Match             65.2%;  Score 349;  DB 6;  Length 89;
  Best Local Similarity   59.1%;  Pred. No. 5.26e-27;
  Matches   39;  Conservative   14;  Mismatches   13;  Indels    0;  Gaps    0;

Db       24 CLVCSDEASGCHYGVLTCGSCKVFFKRAVEGQHNYLCAGRNDCIIDKIRRKNCPACRYRK 83
            | |||| ||| |||| :| :|| ||||:::|:::|:|:: |:| ||| |||:| ||| ||
Qy        1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRLRK 60

Db       84 CLQAGM 89
            | : ||
Qy       61 CYEVGM 66


RESULT    5
ID   US-08-087-151-3      STANDARD;      PRT;    89 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 3, Application US/08087151.
XX
CC    Sequence 3, Application US/08087151
CC    Patent No. 5468624
CC     GENERAL INFORMATION:
CC       APPLICANT:  E. Brad Thompson
CC       APPLICANT:  Lynne V. Nazareth
CC       TITLE OF INVENTION:  ANTI-TUMOR ACTIVITY OF A MODIFIED FRAGMENT
CC       TITLE OF INVENTION:  OF THE GLUCOCORTICOID RECEPTOR
CC       NUMBER OF SEQUENCES:  3
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Arnold, White & Durkee
CC         STREET:  P.O. Box 4433
CC          CITY:  Houston
CC          STATE:  Texas
CC          COUNTRY:  USA
CC          ZIP:  77210
CC        COMPUTER READABLE FORM:
CC          MEDIUM TYPE:  Floppy disk
CC          COMPUTER:  IBM PC compatible
CC          OPERATING SYSTEM:  PC-DOS/MS-DOS
CC          SOFTWARE:  WordPerfect 5.1
CC        CURRENT APPLICATION DATA:
CC          APPLICATION NUMBER:  US/08/087,151
CC          FILING DATE:  July 7, 1993
CC          CLASSIFICATION:  435
CC        ATTORNEY/AGENT INFORMATION:
CC          NAME:  Hodgins, Daniel S.
CC          REGISTRATION NUMBER:  31,026
CC          REFERENCE/DOCKET NUMBER:  UTSB 141/HTD
CC        TELECOMMUNICATION INFORMATION:
CC          TELEPHONE:  512-320-7200
CC          TELEFAX:  512-474-7577
CC          TELEX:  No. 5468624 Applicable
CC      INFORMATION FOR SEQ ID NO:  3.
CC        SEQUENCE CHARACTERISTICS:
CC          LENGTH:  89 amino acid residues
CC          TYPE:  amino acid
CC          STRANDEDNESS:  single
CC          TOPOLOGY:  linear
CC        MOLECULE TYPE:  peptide
SQ   SEQUENCE   89 AA; 9644 MW; 34324 CN;

  Query Match             65.2%;  Score 349;  DB 5;  Length 89;
  Best Local Similarity  59.1%;  Pred. No. 5.26e-27;
  Matches    39;  Conservative   14;  Mismatches  13;  Indels   0;  Gaps   0;

Db        24 CLVCSDEASGCHYGVLTCGSCKVFFKRAVEGQHNYLCAGRNDCIIDKIRRKNCPACRYRK 83
             | |||| ||| |||| :| .|| ||||.. |   | | : | | ||| ||| | ||| ||
Qy         1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCIIDKNRRKSCQACRLRK 60

Db        84 CLQAGM 89
             | : ||
Qy        61 CYEVGM 66


RESULT     6
ID   5223606-7      STANDARD;      PRT;   386 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Patent No. 5223606.
XX
CC   Patent No. 5223606
CC        APPLICANT:  BLAUDIN DE THE, HUGHES;MARCHIO, AGNES;TIOLLAIS,
CC   PIERRE;DEJEAN, ANNE
CC        TITLE OF INVENTION  STEROID/THYROID HORMONE RECEPTOR-RELATED
CC   PROTEIN INAPPPROPRIATELY EXPRESSED IN HUMAN HEPATOCELLULAR CARCINOMA
CC        NUMBER OF SEQUENCES: 11
CC        CURRENT APPLICATION DATA:
CC          APPLICATION NUMBER:  US/07/134,130
CC          FILING DATE: 17-DEC-1987
CC        PRIOR APPLICATION DATA:
CC   SEQ ID NO:7:
CC          LENGTH: 356
SQ   SEQUENCE   386 AA; 44008 MW; 864027 CN;

  Query Match             65.2%;  Score 349;  DB 1;  Length 356;
  Best Local Similarity  59.1%;  Pred. No. 5.26e-27;
  Matches    39;  Conservative   14;  Mismatches  13;  Indels   0;  Gaps   0;

Db         1 CLVCSDEASGCHYGVLTCGSCKVFFKRAVEGQHNYLCAGRNDCIIDKIRRKNCPACRYRK 60
             | |||| ||| |||| :| .|| ||||...|...|.:.. |.| ||| |||:| ||| ||
Qy         1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCIIDKNRRKSCQACRLRK 60
```

```
Db        61 CLQAGM 66
             | : ||
Qy        61 CYEVGM 66


RESULT     7
ID   US-07-716-827C-5      STANDARD;      PRT;   795 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 5, Application US/07716827C.
XX
CC    Sequence 5, Application US/07716827C
CC    Patent No. 5215916
CC     GENERAL INFORMATION:
CC       APPLICANT:  Simons Jr , Stoney S.
CC       APPLICANT:  Yamamoto, K. R.
CC       APPLICANT:  Chakraborti, P. K.
CC       APPLICANT:  Garabedian, M. J.
CC       TITLE OF INVENTION   SUPER GLUCOCORTICOID RECEPTORS
CC       NUMBER OF SEQUENCES:  5
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Cushman, Darby & Cushman
CC         STREET:  Eleventh floor, 1615 L Street, N.W.
CC         CITY:  Washington
CC         STATE:  D.C.
CC         COUNTRY:  U.S.A.
CC         ZIP:  20036-5601
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/716,827C
CC         FILING DATE:  19910619
CC         CLASSIFICATION:  435
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Scott, Watson T.
CC         REGISTRATION NUMBER:  26,581
CC         REFERENCE/DOCKET NUMBER:  WIS/5683/84453
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (202)8613000
CC         TELEFAX:  (202)822-8944
CC         TELEX:  6714627 CUSH
CC     INFORMATION FOR SEQ ID NO.  5:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  795 amino acids
CC         TYPE:  AMINO ACID
CC         TOPOLOGY:  linear
SQ   SEQUENCE   795 AA; 87498 MW; 3326890 CN;

  Query Match             65.2%;  Score 349;  DB 3;  Length 795;
  Best Local Similarity  59.1%;  Pred. No. 5.26e-27;
  Matches    39;  Conservative   14;  Mismatches  13;  Indels   0;  Gaps   0;

Db       440 CLVCSDEASGCHYGVLTCGSCKVFFKRAVEGQHNYLCAGRNDCIIDKIRRKNCPACRYRK 499
             | |||| ||| |||| :| :|| ||||:::|:::|:|:: |:| ||| |||:| ||| ||
Qy         1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCIIDKNRRKSCQACRLRK 60

Db       500 CLQAGM 505
             | : ||
Qy        61 CYEVGM 66


RESULT     8
ID   5260432-2      STANDARD;      PRT;   492 AA.
XX
AC   xxxxxx
XX
```

```
DT   01-JAN-1900
XX
DE   Patent No. 5260432.
XX
CC   Patent No. 5260432
CC        APPLICANT: TAKAKU, FUMIMARO;ISHIKAWA, TAKASHI;IMAWARI, MICHIO;
CC   EVANS, RONALD M.;UMESONO, KAZUHIKO
CC        TITLE OF INVENTION  HUMAN GAMMA RETINOIC ACID RECEPTOR DNA
CC        NUMBER OF SEQUENCES: 4
CC        CURRENT APPLICATION DATA:
CC          APPLICATION NUMBER:  US/07/370,407
CC          FILING DATE: 22-JUN-1989
CC   SEQ ID NO:2:
CC          LENGTH: 454
SQ   SEQUENCE   492 AA; 54527 MW; 1333563 CN;

  Query Match               63.2%;  Score 338;  DB 1.  Length 454;
  Best Local Similarity  60.6%;  Pred  No  7.59e-26;
  Matches    40;  Conservative   10,  Mismatches  16;  Indels   0;  Gaps   0;

Db       90 CFVCNDKSSGYHYGVSSCEGCKGFFRRSIQKNMVYTCHRDKNCIINKVTRNRCQYCRLQK 149
            | ||:| :||||||| |||||:||.||||  :   | |     : | |:|   |: || |||:|
Qy        1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 60

Db      150 CFEVGM 155
            |:||||
Qy       61 CYEVGM 66


RESULT    9
ID   5223606-3      STANDARD;      PRT;    398 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Patent No. 5223606.
XX
CC   Patent No. 5223606
CC        APPLICANT:  BLAUDIN DE THE, HUGHES;MARCHIO, AGNES;TIOLLAIS,
CC   PIERRE;DEJEAN, ANNE
CC        TITLE OF INVENTION: STEROID/THYROID HORMONE RECEPTOR-RELATED
CC   PROTEIN INAPPROPRIATELY EXPRESSED IN HUMAN HEPATOCELLULAR CARCINOMA
CC        NUMBER OF SEQUENCES: 11
CC        CURRENT APPLICATION DATA:
CC          APPLICATION NUMBER:  US/07/134,130
CC          FILING DATE: 17-DEC-1987
CC        PRIOR APPLICATION DATA:
CC   SEQ ID NO:3:
CC          LENGTH: 368
SQ   SEQUENCE   398 AA; 44842 MW; 907111 CN;

  Query Match               62.2%;  Score 333;  DB 1;  Length 368;
  Best Local Similarity  59.1%;  Pred. No. 2.55e-25,
  Matches    39;  Conservative   10,  Mismatches  17;  Indels   0;  Gaps   0;

Db        1 CFVCQDKSSGYHYGVSACEGCKGFFRRSIQKNMIYTCHRDKNCVINKVTRNRCQYCRLQK 60
            | || | :||||||| :|||||:||:||||  :   | |     : | |:|   |: || |||:|
Qy        1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 60

Db       61 CFEVGM 66
            |:||||
Qy       61 CYEVGM 66


RESULT   10
ID   5223606-2      STANDARD;      PRT;    485 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Patent No. 5223606.
XX
CC   Patent No. 5223606
CC        APPLICANT:  BLAUDIN DE THE, HUGHES;MARCHIO, AGNES;TIOLLAIS,
CC   PIERRE;DEJEAN, ANNE
CC        TITLE OF INVENTION  STEROID/THYROID HORMONE RECEPTOR-RELATED
CC   PROTEIN INAPPROPRIATELY EXPRESSED IN HUMAN HEPATOCELLULAR CARCINOMA
CC        NUMBER OF SEQUENCES: 11
CC        CURRENT APPLICATION DATA:
CC          APPLICATION NUMBER   US/07/134,130
CC          FILING DATE: 17-DEC-1987
CC        PRIOR APPLICATION DATA:
CC   SEQ ID NO:2:
CC          LENGTH: 448
SQ   SEQUENCE   485 AA; 54437 MW; 1345912 CN;

  Query Match               62.2%;  Score 333;  DB 1;  Length 448;
  Best Local Similarity  59.1%;  Pred. No. 2.55e-25;
  Matches    39;  Conservative   10;  Mismatches  17;  Indels   0;  Gaps   0;

Db       81 CFVCQDKSSGYHYGVSACEGCKGFFRRSIQKNMIYTCHRDKNCVINKVTRNRCQYCRLQK 140
            | || | :||||||| :|||||:||:||||  :   | |     : | |:|   |: || |||:|
Qy        1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 60

Db      141 CFEVGM 146
            |:||||
Qy       61 CYEVGM 66


RESULT   11
ID   PCT-US92-02320A-4      STANDARD;      PRT;    462 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 4, Application PC/TUS9202320A.
XX
CC    Sequence 4, Application PC/TUS9202320A
CC     GENERAL INFORMATION.
CC       APPLICANT:  Sloan-Kettering Institute, For Cancer Research
CC       TITLE OF INVENTION.  METHODS FOR DETECTION AND TREATMENT OF CANCER
CC       NUMBER OF SEQUENCES:  6
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  COOPER & DUNHAM
CC         STREET:  30 ROCKEFELLER PLAZA
CC         CITY:  NEW YORK
CC         STATE:  NEW YORK
CC         COUNTRY:  U.S.A.
CC         ZIP:  10112
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.24
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER    PCT/US92/02320A
CC         FILING DATE:
CC         CLASSIFICATION:
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 675,084
CC         FILING DATE:  22-MAR-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 673,838
CC         FILING DATE:  22-MAR-1991
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  WHITE, JOHN P
CC         REGISTRATION NUMBER:  28,678
CC         REFERENCE/DOCKET NUMBER:  38694-PCT
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 977-9550
CC         TELEFAX:  (212) 644-0525
CC         TELEX:  (212) 422523 COOP UI
CC     INFORMATION FOR SEQ ID NO:  4:
```

```
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH: 462 amino acids
CC         TYPE: amino acid
CC         TOPOLOGY: linear
CC       MOLECULE TYPE: protein
SQ   SEQUENCE   462 AA; 50771 MW; 1101259 CN;

  Query Match             62.2%;  Score 333;  DB 10;  Length 462;
  Best Local Similarity  59.1%;  Pred. No. 2.55e-25;
  Matches   39;  Conservative   10;  Mismatches  17;  Indels   0;  Gaps   0;

Db      88 CFVCQDKSSGYHYGVSACEGCKGFFRRSIQKNMVYTCHRDKNCIINKVTRNRCQYCRLQK 147
              | |  :||||||| :|||||:||:||||  :   | |     : |  |:|   |: || ||||:|
Qy       1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRLRK 60

Db     148 CFEVGM 153
           |:||||
Qy      61 CYEVGM 66


RESULT   12
ID   PCT-US92 02320A-2     STANDARD;     PRT;    797 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 2, Application PC/TUS9202320A.
XX
CC    Sequence 2, Application PC/TUS9202320A
CC     GENERAL INFORMATION:
CC       APPLICANT:  Sloan-Kettering Institute, For Cancer Research
CC       TITLE OF INVENTION:  METHODS FOR DETECTION AND TREATMENT OF CANCER
CC       NUMBER OF SEQUENCES:  6
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  COOPER & DUNHAM
CC         STREET:  30 ROCKEFELLER PLAZA
CC         CITY:  NEW YORK
CC         STATE:  NEW YORK
CC         COUNTRY:  U.S.A.
CC         ZIP:  10112
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.24
CC       CURRENT APPLICATION DATA
CC         APPLICATION NUMBER:  PCT/US92/02320A
CC         FILING DATE:
CC         CLASSIFICATION:
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 675,084
CC         FILING DATE:  22-MAR-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 673,938
CC         FILING DATE:  22-MAR-1991
CC       ATTORNEY/AGENT INFORMATION
CC         NAME:  WHITE, JOHN P
CC         REGISTRATION NUMBER:  28,678
CC         REFERENCE/DOCKET NUMBER:  38694-PCT
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 977-9550
CC         TELEFAX:  (212) 664-0525
CC         TELEX.  (212) 422523 COOP UI
CC     INFORMATION FOR SEQ ID NO:  2:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  797 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   797 AA; 89281 MW; 3152972 CN;

  Query Match             62.2%;  Score 333;  DB 10;  Length 797;
  Best Local Similarity  59.1%,  Pred. No. 2.55e-25;
  Matches   39,  Conservative   10,  Mismatches  17,  Indels   0,  Gaps   0;

Db     423 CFVCQDKSSGYHYGVSACEGCKGFFRRSIQKNMVYTCHRDKNCIINKVTRNRCQYCRLQK 482
           | || |  :||||||| :|||||:||:||||  :   | |     : |  |:|   |: ||  :|:|
Qy       1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRLRK 60

Db     483 CFEVGM 488
           |.||||
Qy      61 CYEVGM 66


RESULT   13
ID   US-08-333-358-8     STANDARD;     PRT;    440 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 8, Application US/08333358.
XX
CC    Sequence 8, Application US/08333358
CC    Patent No. 5571696
CC     GENERAL INFORMATION:
CC       APPLICANT:  EVANS Ph.D., RONALD M.
CC       APPLICANT:  MANGELSDORF Ph.D., DAVID J.
CC       APPLICANT:  ONG Ms., ESTELITA S.
CC       APPLICANT:  ORO Ph.D., ANTHONY E.
CC       APPLICANT:  BORGMEYER Ph.D., UWE K.
CC       APPLICANT:  GIGUERE Ph.D., VINCENT NMN
CC       APPLICANT:  YAO Mr., TSO-PANG NMN
CC       TITLE OF INVENTION:  NOVEL RECEPTORS
CC       NUMBER OF SEQUENCES:  14
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Pretty, Schroeder, Brueggemann & Clark
CC         STREET:  444 So. Flower St., Suite 2000
CC         CITY:  Los Angeles
CC         STATE:  CA
CC         COUNTRY:  US
CC         ZIP:  90071-2921
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/333,358
CC         FILING DATE:
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/761,068
CC         FILING DATE:
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Reiter Ph.D., Stephen E.
CC         REGISTRATION NUMBER:  31192
CC         REFERENCE/DOCKET NUMBER:  P31 8936
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (619) 535-9001
CC         TELEFAX:  (619) 535-8949
CC     INFORMATION FOR SEQ ID NO:  8:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  440 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   440 AA; 49883 MW; 957063 CN;

  Query Match             59.8%;  Score 320;  DB 6;  Length 440;
  Best Local Similarity  54.5%;  Pred. No. 5.93e-24;
  Matches   36;  Conservative   11;  Mismatches  19;  Indels   0;  Gaps   0;

Db      91 CSVCGDKASGFHYNVLSCEGCKGFFRRSVIKGAHYICHSGGHCPMDTYMRRKCQECRLRK 150
           |:|||:| ||||:|| | ||||||:||:||:      ||| :  :| :|   |: || |||||
Qy       1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRLRK 60

Db     151 CRQAGM 156
           | : ||
Qy      61 CYEVGM 66


RESULT   14
ID   5171671 2      STANDARD;      PRT;    501 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Patent No. 5171671.
XX
CC   Patent No. 5171671
CC         APPLICANT: EVANS, RONALD M.;ONG, ESTELITA S.;SEGUI,
CC   PRUDIMAR S.;THOMPSON, CATHERINE C.;UEMSONO, KAZUHIKO
CC   GUIGUERE, VINCENT
CC         TITLE OF INVENTION: RETINOIC ACID RECEPTOR COMPOSITION
CC         NUMBER OF SEQUENCES: 2
CC         CURRENT APPLICATION DATA:
CC           APPLICATION NUMBER: US/07/546,256
CC           FILING DATE: 06-AUG-1990
CC         PRIOR APPLICATION DATA:
CC           APPLICATION NUMBER: 276,536
CC           FILING DATE: 30-NOV-1988
CC           APPLICATION NUMBER: 128,331
CC           FILING DATE: 02-DEC-1987
CC   SEQ ID NO:2:
CC         LENGTH: 462
SQ   SEQUENCE   501 AA, 55050 MW, 1419525 CN.

  Query Match             59.3%;  Score 317;  DB 1;  Length 462;
  Best Local Similarity   57.6%;  Pred. No. 1.22e-23;
  Matches   38;  Conservative   10;  Mismatches   18;  Indels   0;  Gaps   0;

Db      88 CFVCQDKSSGTHYGVSACEGCKGFFRRSIQKNMVYTCHRDKNCIINKVTRNRCQYCRLQK 147
           | || | :|| |||| :|||||:||:||||  :  | |    : | |:|  |: || |||:|
Qy       1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRLRK 60

Db     148 CFEVGM 153
           |:||||
Qy      61 CYEVGM 66


RESULT   15
ID   PCT US95-16311-2     STANDARD;      PRT;    451 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 2, Application PC/TUS9516311.
XX
CC    Sequence 2, Application PC/TUS9516311
CC     GENERAL INFORMATION:
CC         APPLICANT:  Moore, David
CC         APPLICANT:  Seol, Wongi
CC         APPLICANT:  Choi, Hueng-Sik
CC         TITLE OF INVENTION:  RETINOID X RECEPTOR-INTERACTING
CC         TITLE OF INVENTION:  POLYPEPTIDES AND RELATED MOLECULES AND METHODS
CC         NUMBER OF SEQUENCES:  17
CC         CORRESPONDENCE ADDRESS:
CC           ADDRESSEE:  Fish & Richardson P.C.
CC           STREET:  225 Franklin Street, Suite 3100
CC           CITY:  Boston
CC           STATE:  MA
CC           COUNTRY:  USA
CC           ZIP:  02110-2804
CC         COMPUTER READABLE FORM:
CC           MEDIUM TYPE:  Floppy disk
CC           COMPUTER:  IBM PC compatible
CC           OPERATING SYSTEM:  PC-DOS/MS-DOS
CC           SOFTWARE:  PatentIn Release #1.0, Version #1.30
CC         CURRENT APPLICATION DATA:
CC           APPLICATION NUMBER:  PCT/US95/16311
CC           FILING DATE:
CC         PRIOR APPLICATION DATA:
CC           APPLICATION NUMBER:  08/372,652
CC           FILING DATE:  13-JAN-1995
CC         ATTORNEY/AGENT INFORMATION:
CC           NAME:  Clark, Paul T.
CC           REGISTRATION NUMBER:  30,162
CC           REFERENCE/DOCKET NUMBER:  00786/246001
CC         TELECOMMUNICATION INFORMATION:
CC           TELEPHONE:  617/542-5070
CC           TELEFAX:  617/542-8906
CC           TELEX:  200154
CC       INFORMATION FOR SEQ ID NO:  2:
CC         SEQUENCE CHARACTERISTICS:
CC           LENGTH:  451 amino acids
CC           TYPE:  amino acid
CC           STRANDEDNESS:  not relevant
CC           TOPOLOGY:  linear
CC         MOLECULE TYPE.  protein
SQ   SEQUENCE   451 AA; 51979 MW; 1018265 CN;

  Query Match             57.9%;  Score 310;  DB 13;  Length 451;
  Best Local Similarity   54.5%;  Pred. No. 6.63e 23;
  Matches   36;  Conservative    9;  Mismatches   21;  Indels   0;  Gaps   0;

Db     101 CVVCGDRASGYHYNALICEGCKGFFRRSIIKNAVYKCKNGGNCVMDMYMRRKCQECRLRK 160
           | ||:| ||||||   :|||||:||:|||  :  | |      | :|:  |: || |||||
Qy       1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRLRK 60

Db     161 CREMGM 166
           | |:||
Qy      61 CYEVGM 66


Search completed: Thu Dec 18 08:29:53 1997
Job time : 7 secs.
```

```
**********************************************************************
   __  __   ____    ____   ____    ____  _   _
  |  \/  | |  _ \  / ___| |  _ \  / ___|| | | |
  | |\/| | | |_) | \___ \ | |_) || |    | |_| |
  | |  | | |  __/   ___) ||  _ < | |___ |  _  |
  |_|  |_| |_|     |____/ |_| \_\ \____||_| |_|  (TM)
**********************************************************************

         Release 2.1D John F. Collins, Biocomputing Research Unit.
         Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
                Distribution rights by IntelliGenetics, Inc.

MPsrch_nn   n.a. - n.a. database search, using Smith-Waterman algorithm

Run on:          Thu Dec 18 21:09:45 1997; MasPar time 521.37 Seconds
                                          1000.431 Million cell updates/sec
Tabular output not generated

Title:           >US-08-826-361A-24
Description:     (1-1898) from US08826361A.seq
Perfect Score:   1898
N.A. Sequence:       1 CAGGAATCTTTGAGAACATT..........TGAATGGAGGTTTTCCTACAC 1898
     Comp                GTCCTTAGAAACTCTTGTAA..........ACTTACCTCCAAAAGGATGTG

Scoring table:   TABLE default
                 Gap 6

Nmatch   STD :   Dbase 0; Query 0

Searched:        359097 seqs, 137407154 bases x 2

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database:        EST-STS-THREE
                   1:EST199 2:EST200 3:EST201 4:EST202 5:EST203 6:EST204
                   7:EST205 8:EST206 9:EST207 10:EST208 11:EST209 12:EST210
                   13:EST211 14:EST212 15:EST213 16:EST214 17:EST215
                   18:EST216 19:EST217 20:EST218 21:EST219 22:EST220
                   23:EST221 24:EST222 25:EST223 26:EST224 27:EST225
                   28:EST226 29:EST227 30:EST228 31:EST229 32:EST230
                   33:EST231 34:EST232 35:EST233 36:STS1 37:STS2 38:STS3
                   39:STS4 40:STS5 41:STS6 42:STS7 43:STS8 44:STS9 45:STS10
                   46:STS11 47:STS12 48:STS13
Database:        EST-STS-FOUR
                   49:gnEST1 50:gnEST2 51:gnEST3 52:gnEST4 53:gnEST5
                   54:gnEST6 55:gnEST7 56:gnEST8 57:gnEST9 58:gnEST10
                   59:gnEST11 60:gnEST12 61:gnEST13 62:gnEST14 63:gnEST15
                   64:gnEST16 65:gnEST17 66:gnEST18 67:gnEST19 68:gnEST20
                   69:gnEST21 70:gnEST22 71:gnEST23 72:gnEST24 73:gnEST25
                   74:gnSTS 75:enEST1 76:enEST2 77:enEST3 78:enEST4
                   79:enEST5 80:enEST6 81:enEST7 82:enEST8 83:enEST9
                   84:enEST10 85:enEST11 86:enEST12 87:enEST13 88:enEST14
                   89:enEST15 90:enEST16 91:enEST17 92:enEST18 93:enEST19
                   94:enEST20 95:enEST21 96:enEST22 97:enEST23 98:enEST24
                   99:enEST25 100:enEST26 101:enEST27 102:enEST28
                   103:enEST29 104:enEST30 105:enEST31 106:enSTS 107:ueEST1
                   108:ueEST2

Statistics:      Mean 11.647;  Variance 2.000;  scale 5.824

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                               SUMMARIES

                   %
Result          Query
```

```
     No.  Score  Match Length DB  ID            Description           Pred. No.
   ---------------------------------------------------------------------------
c     1    158    8.3    531 81  HS1143676     zr14e02.r1 Stratagene  1.27e-290
c     2    158    8.3    531 61  AA224131      zr14e02.r1 Stratagene  1.27e-290
      3     53    2.8   1212 44  G29564        human STS SHGC-35295.  1.09e-56
c     4     52    2.7    390 96  MM7156        ma33b04.r1 Soares mou  1.03e-54
c     5     40    2.1    593 28  AA023625      mh80b06.r1 Soares mou  6.13e-32
      6     26    1.4    280 25  N59515        yw57g05.r1 Soares fet  6.25e-09
      7     24    1.3    156 36  DM36D7S       D. melanogaster STS d  3.63e-06
      8     24    1.3    526 97  MM94013       mb96d08.r1 Soares mou  3.63e-06
      9     22    1.2    250 27  I93798        lmEST0161 LmLV39cDNA   1.34e-03
c    10     22    1.2    326 85  HS1164487     111371.seq.F Human fe  1.34e-03
c    11     22    1.2    326 69  AA249798      111371.seq.F Human fe  1.34e-03
     12     22    1.2    350 27  I67398        lmES10039 LmLV39cDNA   1.34e-03
     13     23    1.2    358 65  AA234422      zr72d02.r1 Soares NhH  7.42e-05
     14     23    1.2    358 93  HS1152547     zr72d02.r1 Soares NhH  7.42e-05
     15     22    1.2    382 30  AA057786      zf15e04.s1 Soares fet  1.34e-03
c    16     22    1.2    399 21  AA021604      ze67c04.s1 Soares fet  1.34e-03
     17     22    1.2    424 89  HSW0986       ze08h09.r1 Soares fet  1.34e-03
     18     22    1.2    424  2  W35098        ze08h09.r1 Soares fet  1.34e-03
     19     22    1.2    436 104  MMAA85071     mv85g09.r1 GuayWoodfo  1.34e-03
     20     22    1.2    436 20  AA208507      mv85g09.r1 Beier mous  1.34e-03
     21     22    1.2    436 70  AA208507      mv85g09.r1 GuayWoodfo  1.34e-03
     22     22    1.2    452 23  AA024507      zu76a08.s1 Soares fet  1.34e-03
c    23     23    1.2    458 10  AA195198      zp95e10.s1 Stratagene  7.42e-05
c    24     22    1.2    498 11  AA179181      zp46b03.r1 Stratagene  1.34e-03
c    25     22    1.2    501 84  HS1151180     zr69e09.r1 Soares NhH  1.34e-03
c    26     22    1.2    501 65  AA233250      zr69e09.r1 Soares NhH  1.34e-03
c    27     22    1.2    545 19  AA202294      LD02403 5prime LD [illegible]  1.34e-03
     28     23    1.2   1145 38  G05379        human STS WI-7058.     7.42e-05
     29     21    1.1    285 26  SSC94056      S.scrofa mRNA; expres  2.13e-02
     30     21    1.1    320  2  W95560        ze02b04.r1 Soares fet  2.13e-02
c    31     21    1.1    359 36  G00807        fruit fly STS Dm0760   2.13e-02
     32     21    1.1    384 65  AA235237      zs38g05.r1 Soares NhH  2.13e-02
     33     21    1.1    384 83  HS1153545     zs38g05.r1 Soares NhH  2.13e-02
c    34     21    1.1    394  4  AA114039      zm29e11.r1 Stratagene  2.13e-02
c    35     21    1.1    400 40  G11391        human STS SHGC-10167.  2.13e-02
     36     21    1.1    438 42  G24207        human STS WI-12034.    2.13e-02
c    37     21    1.1    443 72  AA253566      CpEST.158 pSKIIminusC  2.13e-02
c    38     21    1.1    462 90  HSW6574       ze04h01.r1 Soares fet  2.13e-02
c    39     21    1.1    462  2  W32657        ze04h01.r1 Soares fet  2.13e-02
c    40     21    1.1    465 55  AA146657      mo03h09.r1 Stratagene  2.13e-02
c    41     21    1.1    498 55  AA174380      mo96a11.r1 Stratagene  2.13e-02
c    42     21    1.1    528 108  MM1139113     mv68c05.r1 Soares mou  2.13e-02
c    43     21    1.1    528 43  AA220665      mv68c05.r1 Soares mou  2.13e-02
c    44     21    1.1    555 104  MMAA81229     mw71g11.r1 Soares mou  2.13e-02
c    45     21    1.1    557 17  AA193507      mt54g08.r1 Stratagene  2.13e-02


                              ALIGNMENTS

RESULT     1
ID    HS1143676  standard; RNA; EST; 531 BP.
AC    AA224131;
NI    g1844690
DT    22-FEB-1997 (Rel. 51, Created)
DT    22-FEB-1997 (Rel. 51, Last updated, Version 1)
DE    zr14e02.r1 Stratagene NT2 neuron (#937233) Homo sapiens cDNA clone
DE    648794 5'.
KW    EST.
OS    Homo sapiens (human)
OC    Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
OC    Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN    [1]
RP    1-531
RA    Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,
RA    Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,
RA    Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,
RA    Trevaskis E., Waterston R., Williamson A., Wohldmann P., Wilson R.;
RT    "The WashU-Merck EST Project";
RL    Unpublished.
CC    Contact: Wilson RK WashU-Merck EST Project Washington University
CC    School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis,
CC   MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:
CC   est@watson.wustl.edu This clone is available royalty-free through
CC   LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for
CC   further information. Seq primer: -28m13 rev1 ET from Amersham High
CC   quality sequence stop: 357.
FH   Key             Location/Qualifiers
FH
FT   source          1..531
FT                   /organism="Homo sapiens"
FT                   /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
FT                   XhoI; Cloned unidirectionally. Primer: Oligo dT.
FT                   Differentiated, post mitotic hNT neurons. Average insert
FT                   size: 1.5 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5'
FT                   GAATTCGGCACGAG 3' -3' adaptor sequence: 5'
FT                   CTCGAGTTTTTTTTTTTTTTTTTT 3'"
FT                   /clone="648794"
FT                   /clone_lib="Stratagene hNT neuron (#937233)"
FT                   /dev_stage="hNT neurons"
FT                   /lab_host="SOLR (kanamycin resistant)"
FT   mRNA            <1..>531
SQ   Sequence 531 BP; 192 A; 74 C; 112 G; 146 T; 7 other;

  Query Match             8.3%;  Score 158;  DB 81;  Length 531;
  Best Local Similarity 100.0%;  Pred. No. 1.27e-290;
  Matches  158;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Db       27 cgtgctccaggggtaagatggattgactgcagttgtaggaggaaggagaattaaggctag 86
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Cp      158 CGTGCTCCAGGGGTAAGATGGATTGACTGCAGTTGTAGGAGGAAGGAGAATTAAGGCTAG 99

Db       87 atggtgagttttttatatccatgtcttgagataacagctgagaaaacaccttgcaagaag 146
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Cp       98 ATGGTGAGTTTTTTATATCCATGTCTTGAGATAACAGCTGAGAAAACACCTTGCAAGAAG 39

Db      147 aggcacaaaggtcattataatgttctcaaagattcgtg 184
            ||||||||||||||||||||||||||||||||||||||
Cp       38 AGGCACAAAGGTCATTATAATGTTCTCAAAGATTCGTG 1


RESULT      2
LOCUS       AA224131      531 bp    mRNA            EST       19-FEB-1997
DEFINITION  zr14e02.r1 Stratagene hNT neuron (#937233) Homo sapiens cDNA clone
            648794 5'.
ACCESSION   AA224131
NID         g1844690
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 531)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E.,
            Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -28m13 rev1 ET from Amersham
            High quality sequence stop: 357.
FEATURES             Location/Qualifiers
     source          1..531
                     /organism="Homo sapiens"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI. Cloned unidirectionally. Primer: Oligo dT.
                     Differentiated, post mitotic hNT neurons. Average insert
                     size: 1.5 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5'
                     GAATTCGGCACGAG 3' -3' adaptor sequence: 5'
                     CTCGAGTTTTTTTTTTTTTTTTTT 3'"
                     /clone="648794"
                     /clone_lib="Stratagene hNT neuron (#937233)"
                     /dev_stage="hNT neurons"
                     /lab_host="SOLR (kanamycin resistant)"
     mRNA            <1..>531
BASE COUNT      192 a     74 c    112 g    146 t      7 others
ORIGIN

  Query Match             8.3%;  Score 158;  DB 62;  Length 531;
  Best Local Similarity 100.0%;  Pred. No. 1.27e-290;
  Matches  158;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Db       27 cgtgctccaggggtaagatggattgactgcagttgtaggaggaaggagaattaaggctag 86
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Cp      158 CGTGCTCCAGGGGTAAGATGGATTGACTGCAGTTGTAGGAGGAAGGAGAATTAAGGCTAG 99

Db       87 atggtgagttttttatatccatgtcttgagataacagctgagaaaacaccttgcaagaag 146
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Cp       98 ATGGTGAGTTTTTTATATCCATGTCTTGAGATAACAGCTGAGAAAACACCTTGCAAGAAG 39

Db      147 aggcacaaaggtcattataatgttctcaaagattcgtg 184
            ||||||||||||||||||||||||||||||||||||||
Cp       38 AGGCACAAAGGTCATTATAATGTTCTCAAAGATTCGTG 1


RESULT      3
LOCUS       G28564       2202 bp    DNA             STS       01-JUL-1996
DEFINITION  human STS SHGC-35295.
ACCESSION   G28564
NID         g1408379
KEYWORDS    STS sequence; primer; sequence tagged site.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata, Eutheria, Primates, Catarrhini, Hominidae, Homo.
REFERENCE   1  (bases 1 to 2202)
  AUTHORS   Myers,R.M.
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Richard M. Myers
            Stanford Human Genome Center (SHGC)
            Stanford University School of Medicine
            Department of Genetics, M-344, Stanford, CA 94305, USA
            Tel: 4157253687
            Fax: 4157259689
            Email: myers@shgc.stanford.edu

            Primer A: AATTGAACTCACTGCTTTTGAGG
            Primer B: CAGCTTCAGGTATCATCTTCACC
            STS size: 206
            PCR Profile:
                   Initial incubation: 94 degrees C for 90 seconds

                   Denaturation:       94 degrees C for 15 seconds
                   Annealing:          62 degrees C for 23 seconds
                   Polymerization:     72 degrees C for 30 seconds
                   PCR Cycles:         30
                   Thermal Cycler:     Perkin Elmer 9600
            Protocol:
                   Template:        25 ng
                   Primer:          each 1 uM
                   dNTPs:           each 200 uM
                   Taq Polymerase:  0.05 units/ul
                   Total Vol:       10 ul

            Buffer:
                   MgCl2:           2.5 mM
                   KCl:             50  mM
```

```
                  Tris-HCl:        20  mM
                  pH:              8.3

           Prepared with primer pairs provided by Sandon, derived from M23960
           -- Washington University/Merck EST sequence
FEATURES             Location/Qualifiers
     source          1..2202
                     /organism="Homo sapiens"
                     /map="12"
     STS             1926..2131
     primer_bind     1926..1948
     primer_bind     complement(2109..2131)
BASE COUNT      720 a    442 c    439 g    601 t
ORIGIN

  Query Match             2.8%;  Score 53;  DB 44;  Length 2202;
  Best Local Similarity  61.9%;  Pred. No. 1.09e-56;
  Matches  138;  Conservative    0;  Mismatches   85;  Indels    0;  Gaps    0;

Db     368 aggaccaaataagtttttgatctttgcgtagtatgtggagacaaagcatcaggacgtca 427
            |||  |||| | || |   | |   | ||||   || ||   | ||   | ||||| |||   |||
Qy     496 AGGTTCAAAGAGGATGCTCACTTCTGCGCTGTCTGGCAGCGATTAGGCATCGGGATATCA 555

Db     428 ttatggagcagtaacttgtgaaggctgcaaaggatttttaaaagaagcatccgaaaaaa 487
             |||||||       | |||||||| || || |  |||||||||||||||| | |  | |
Qy     556 CTATGGAGTTTTGTGGTGTGAAGGATGTAAGGCTTTTTTTAAAAGAAGCATTTAAGGACA 615

Db     488 tttagtatattcatgtcgaggatcaaaggattgtattattaataagcaccacgaaacg 547
           |   |  |||    |||| ||    |||| | ||||  ||  ||||  |||  || ||  |
Qy     616 TAATGATTATATTTGTCCAGCTACAAATCAGTGTACAATCGATAAAAACCGGCGCAAGAG 675

Db     548 ctgtcaatactgcaggttacagagatgtattgcgtttggaatg 590
           ||| ||    |||| |  | | ||  |||    |   | ||||||
Qy     676 CTGCCAGGCCTGCCGACTTTGGAAGTGTTACGAAGTGGGAATG 718


RESULT    4
ID   MM7156     standard; RNA; EST; 390 BP.
AC   W13715;
NI   g1287753
DT   30-APR-1996 (Rel. 47, Created)
DT   04-MAR-1997 (Rel. 51, Last updated, Version 2)
DE   ma83b04.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone 317263 5'
DE   similar to gb:L13740 EARLY RESPONSE PROTEIN NAK1 (HUMAN);.
KW   EST.
OS   Mus musculus (house mouse)
OC   Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
OC   Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
OC   Mus.
RN   [1]
RP   1-390
RA   Marra M., Hillier L., Allen M., Bowles M., Dietrich N.,
RA   Dubuque T., Geisel S., Kucaba T., Lacy M., Le M., Martin J.,
RA   Morris M., Schellenberg K., Steptoe M., Tan F., Underwood K.,
RA   Moore B., Theising B., Wylie T., Lennon G., Soares B., Wilson R.,
RA   Waterston R.;
RT   "The WashU-HHMI Mouse EST Project";
RL   Unpublished.
CC   Contact: Marra M/Mouse EST Project WashU-HHMI Mouse EST Project
CC   Washington University School of MedicineP 4444 Forest Park Parkway,
CC   Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810
CC   Email: mouseest@watson.wustl.edu This clone is available
CC   royalty-free through LLNL ; contact the IMAGE Consortium
CC   (info@image.llnl.gov) for further information. Seq primer: ETPrimer
CC   High quality sequence stop: 385.
FH   Key             Location/Qualifiers
FH
FT   source          1..390
FT                   /organism="Mus musculus"
FT                   /note="Vector: pT7T3D (Pharmacia) with a modified
FT                   polylinker. Site_1: Not I. Site_2: Eco RI. 1st strand cDNA
FT                   was primed with a Not I - oligo(dT) primer [5'
FT                   TGTTACCAATCTGAAGTGGGAGCGGCCGCATTTTTTTTTTTTTTTTTTT 3'],
FT                   double-stranded cDNA was size selected, ligated to Eco RI
FT                   adapters (Pharmacia), digested with Not I and cloned into
FT                   the Not I and Eco RI sites of a modified pT7T3 vector
FT                   (Pharmacia). Library went through one round of
FT                   normalization to a Cot = 5. Library constructed by Bento
FT                   Soares and M Fatima Bonaldo. RNA was kindly provided by Dr
FT                   Minoru Ko (Wayne State University)."
FT                   /clone="317263"
FT                   /clone_lib="Soares mouse p3NMF19.5"
FT                   /dev_stage="19.5 dpc total fetus"
FT                   /lab_host="DH10B (ampicillin resistant)"
FT   mRNA            <1..>390
SQ   Sequence 390 BP; 70 A; 106 C; 114 G; 100 T; 0 other;

  Query Match             2.7%;  Score 52;  DB 36;  Length 390;
  Best Local Similarity  62.7%;  Pred. No. 1.03e-54;
  Matches  128;  Conservative    0;  Mismatches   76;  Indels    0;  Gaps    0;

Db      80 cttaaccatcccgacactgagacacttctgaaatctgcagtactgacatcggtttcgccg 139
           ||| ||||| || ||    |  ||||||| ||| || ||||  ||| || |   || || ||
Qy     724 CTTCACCATTCCCACTTCGTAACACTTCCGAAGTCGGCAGGCCTGGCAGCTCTTGCGCCG 665

Db     140 tctcttgtccactgggcagtttttatttgccaggcaaacatattttgcatttttctgcac 199
              | || || | ||  || |   ||   | ||   | |||| ||| |     || |     | |
Qy     664 GTTTTTATTGATTGTACACTGATTTGTAGCTGGACAAATATAATCATTATGTCCTTGAAT 605

Db     200 cgttctcttgaagaagcccttgcagccctcgcaggtgcggactccgtagtgctggcaggc 259
              |||| || || ||| |||| || || || || |   | |||||| ||||| |   |   |
Qy     604 GCTTCTTTTAAAAAAGGCCTTACATCCTTCACACGACCAGACTCCATAGTGATATCCCGA 545

Db     260 agcattgtcgccgcacacggcaca 283
             ||||   |||| ||| || || ||
Qy     544 TGCGTAATCGCTGCAGACACCGCA 521


RESULT    5
LOCUS       AA023625      593 bp    mRNA            EST       21-JAN-1997
DEFINITION  mh80b06.r1 Soares mouse placenta 4NbMP13.5 14.5 Mus musculus cDNA
            clone 457235 5' similar to gb:M12674 ESTROGEN RECEPTOR (HUMAN);
            gb:M38651 Mouse estrogen receptor mRNA, complete cds (MOUSE);.
ACCESSION   AA023625
NID         g1487542
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 593)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:274123
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 492.
FEATURES             Location/Qualifiers
     source          1..593
                     /organism="Mus musculus"
```

```
                    /strain="C57BL/6J"
                    /note="Vector  pT7T3D-Pac (Pharmacia) with a modified
                    polylinker: Site_1: Not I; Site_2  Eco RI. 1st strand cDNA
                    was primed with a Not I  oligo(dT) primer [5'
                    TGTTACCAATCTGAAGTGGGAGCGGCCGCGGAAATTTTTTTTTTTTTTTTTTTTT
                    T 3']; double-stranded cDNA was ligated to Eco RI adaptors
                    (Pharmacia), digested with Not I and cloned into the Not I
                    and Eco RI sites of the modified pT7T3 vector  Library
                    went through one round of normalization, and was
                    constructed by Bento Soares and M.Fatima Bonaldo."
                    /clone="457235"
                    /clone_lib="Soares mouse placenta 4NbMP13.5 14.5"
                    /sex="unknown"
                    /dev_stage="adult"
                    /lab_host="DH10B"
     mRNA           <1..>593
BASE COUNT      144 a    165 c    119 g    165 t
ORIGIN

  Query Match             2.1%;  Score 40;  DB 28;  Length 593;
  Best Local Similarity  74.4%;  Pred. No. 6.13e-32;
  Matches    61;  Conservative    0;  Mismatches  21   Indels    0;  Gaps    0;

Db     174 tctttgcccagttgatcatatgtacccagttcctatctgctggttggtcaataagccca 233
            |||| |||||| ||||||| || |||| |||| | || || |  |||||||   |   ||
Cp    1019 TCTTGGCCCAGCTGATCATGTGTACCAACTCCTTGTCGGCCAACTTGGTCAGGGACATCA 960

Db     234 tcattgaggcttcactgaaggg 255
            |||| ||||| ||  |||||||
Cp     959 TCATGGAGGCCTCGGTGAAGGG 938


RESULT      6
LOCUS       N59515          280 bp    mRNA              EST       28-JAN-1997
DEFINITION  yv57g05.r1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone
            246872 5' similar to SW:LHR1_MOUSE P45448 LIVER RECEPTOR HOMOLOG ;.
ACCESSION   N59515
NID         g1203405
KEYWORDS    EST
SOURCE      human
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 280)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E.,
            Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Insert Length: 728   Std Error: 0.00
            Seq primer: T7
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1..280
                     /organism="Homo sapiens"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Pac I - oligo(dT) primer
                     [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
```

```
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector.  Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
                     /clone="246872"
                     /clone_lib="Soares fetal liver spleen 1NFLS"
                     /sex="male"
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
     mRNA            <1..>280
BASE COUNT       80 a     45 c     75 g     61 t     19 others
ORIGIN

  Query Match             1.4%;  Score 26;  DB 25;  Length 280;
  Best Local Similarity  54.0%;  Pred. No. 6.25e-09;
  Matches    94;  Conservative    0;  Mismatches  80;  Indels    0;  Gaps    0;

Db      26 tcngggtancattangggctcctcanctgtgnaagctgnaaggattttttaagcgnaca 85
            || || || || || ||   ||       |||| | | || ||||  |||||||   | |
Qy     545 TCGGGATATCACTATGGACTCTGGTCGTGTGAAGGATGTAAGGCCTTTTTTAAAAGAAGC 604

Db      86 gtccaaaataataaaagtacactgnatataaacctagnactgccgaattgncaaaacn 145
             | |||     ||||     || |  ||     |    |    |  | ||    ||| |  ||||
Qy     605 ATTCAAGCACATAATGATTATATTTGTCCAGCTACAAATCAGTGTACAATCGATAAAAAC 664

Db     146 cagagaaagcgttgtccttactgtcgttttcaaaaatgtctaagtgttggaatg 199
            | | | ||| | || |     ||| ||  |||  || |||       || ||||||
Qy     665 CGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTACGAAGTGGGAATG 718


RESULT      7
LOCUS       DM36D7S         156 bp    DNA               STS       16-APR-1996
DEFINITION  D. melanogaster STS determined from European Mapping Project
            cosmid.
ACCESSION   Z70880
NID         g1263754
KEYWORDS    sequence tagged site.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Arthropoda;
            Tracheata; Insecta; Pterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 156)
  AUTHORS   European Drosophila Mapping,Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-APR-1996) Michael Ashburner, Department of Genetics,
            Downing St., Cambridge CB2 3EH, England
COMMENT     STS_name = Dm36D7S
            clone_name = 36D7
            STS_from_promoter = SP6
            vector_class = cosmid, Lorist 6
            origin_of_clone = Oregon-R
            in_situ_site_primary = 98C
            BLAST_program = BLASTN
            database_searched = EMBL
            database_version = 45.0 and updates till date_of_search
            date_of_search = 08-01-1996
            BLAST_program = BLASTX
            database_searched = SWISSPROT
            database_version = 32.0
            date_of_search = 15-12-1995.
FEATURES             Location/Qualifiers
     source          1..156
                     /organism="Drosophila melanogaster"
                     /strain="Oregon-R"
                     /clone="36D7"
BASE COUNT       28 a     37 c     30 g     27 t     34 others
ORIGIN

  Query Match             1.3%;  Score 24;  DB 36;  Length 156;
  Best Local Similarity  19.6%;  Pred. No. 3.63e-06;
  Matches     9;  Conservative   26;  Mismatches  11;  Indels    0;  Gaps    0;
```

```
Db      21 gskssrrgrrrgrkrrgagsksskarrgkkcrkskrgssggagtrr 66
           |: :::.| .||....  |........ . | :. : .. |:|
Qy     832 GCCCGAGTGCGGGAGCTGCTGCTGGACGCCCTGAGCCCCGAGCAG 877


RESULT     8
ID   MM94013     standard; RNA; EST; 536 BP.
AC   W20940;
NI   g1297857
DT   08-MAY-1996 (Rel  47  Created)
DT   05-MAR-1997 (Rel. 51, Last updated, Version 2)
DE   mb96d08 r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone 337263 5'
DE   similar to gb:M64497 APOLIPOPROTEIN AI REGULATORY PROTEIN-1
DE   (HUMAN); gb:U07635 Mus musculus ovalbumin upstream promoter
DE   transcription factor (MOUSE),.
KW   EST.
OS   Mus musculus (house mouse)
OC   Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
OC   Vertebrata; Eutheria  Rodentia  Sciurognathi  Muridae  Murinae
OC   Mus.
RN   [1]
RP   1-536
RA   Marra M., Hillier L   Allen M , Bowles M   Dietrich N.,
RA   Dubuque T., Geisel S., Kucaba T., Lacy M., Le M., Martin J.,
RA   Morris M   Schellenberg K , Steptoe M., Tan F., Underwood K.,
RA   Moore B., Theising B., Wylie T , Lennon G , Soares B , Wilson R.,
RA   Waterston R.;
RT   "The WashU-HHMI Mouse EST Project";
RL   Unpublished.
CC   Contact  Marra M/Mouse EST Project WashU-HHMI Mouse EST Project
CC   Washington University School of MedicineP 4444 Forest Park Parkway,
CC   Box 8501  St  Louis  MO 63108 Tel  314 286 1800 Fax  314 286 1810
CC   Email: mouseest@watson.wustl.edu This clone is available
CC   royalty-free through LLNL ; contact the IMAGE Consortium
CC   (info@image llnl gov) for further information. MGI:219663 Seq
CC   primer  FTPrimer High quality sequence stop: 385.
FH   Key             Location/Qualifiers
FH
FT   source          1..536
FT                   /organism="Mus musculus"
FT                   /note="Vector  pT7T3D (Pharmacia) with a modified
FT                   polylinker  Site_1  Not I  Site_2  Eco RI  1st strand cDNA
FT                   was primed with a Not I - oligo(dT) primer [5'
FT                   TGTTACCAATCTGAAGTGGGAGCGGCCGCATTTTTTTTTTTTTTTTTTT 3'],
FT                   double-stranded cDNA was size selected, ligated to Eco RI
FT                   adapters (Pharmacia), digested with Not I and cloned into
FT                   the Not I and Eco RI sites of a modified pT7T3 vector
FT                   (Pharmacia). Library went through one round of
FT                   normalization to a Cot = 5. Library constructed by Bento
FT                   Soares and M Fatima Bonaldo. RNA was kindly provided by Dr

FT                   Minoru Ko (Wayne State University)."
FT                   /clone="337263"
FT                   /clone_lib="Soares mouse p3NMF19 5"
FT                   /dev_stage="19 5 dpc total fetus"
FT                   /lab_host="DH10B (ampicillin resistant)"
FT   mRNA            <1..>536
SQ   Sequence 536 BP; 113 A; 175 C; 164 G; 84 T; 0 other;

  Query Match             1 3%   Score 24   DB 27   Length 536
  Best Local Similarity  66.7%.  Pred  No  3 63e-06,
  Matches    48   Conservative    0   Mismatches  24   Indels    0   Gaps    0;

Db     205 tgtcccatcgaccagcaccaccgcaaccagtgccagtactgcagcctcaaaaagtgcctc 264
           ||| | ||||| : ||: ||:|| ||||||| |||||| || |||||  |
Qy     647 TGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTAC 706

Db     265 aaagtgggcatg 276
            ||||||| |||
Qy     707 GAAGTGGGAATG 718


RESULT     9
```

```
LOCUS       T93398         250 bp    mRNA            EST       23-JAN-1997
DEFINITION  lmEST0161 LmLV39cDNA Leishmania major cDNA clone Lm211 5' END.
ACCESSION   T93398
NID         g726571
KEYWORDS    EST.
SOURCE      Leishmania major.
  ORGANISM  Leishmania major
            Eukaryotae, mitochondrial eukaryotes; Euglenozoa, Kinetoplastida;
            Trypanosomatidae; Leishmania.
REFERENCE   1  (bases 1 to 250)
  AUTHORS   Levick,M P , Blackwell,J.M., Connor,V., Coulson,R.M., Miles,A.,
            Smith,H.E., Wan,K.L. and Ajioka,J.W.
  TITLE     An expressed sequence tag analysis of a full-length, spliced leader
            cDNA library from Leishmania major promastigotes
  JOURNAL   Mol. Biochem. Parasitol. 76 (1-2), 345-348 (1996)
  MEDLINE   97077459
COMMENT
            Contact: Ajioka JW
            Laboratory for Parasite Genome Analysis
            Cambridge University
            Department of Pathology, Tennis Court Road, Cambridge CB2 1QP, UK.
            Tel: 01223333923
            Fax: 01223333923
            Email: jajioka@hgmp.mrc.ac.uk
            Seq primer: T3.
FEATURES             Location/Qualifiers
     source          1..250
                     /organism="Leishmania major"
                     /strain="LV39"
                     /note="Vector, Lambda UNIZAP, Site_1. NotI, Site_2: XhoI;
                     Mid or late log promastigotes.Full length , splice leader
                     oligo used for second strand synthesis, directional
                     cloning NotI at 5' end, XhoI at 3' end."
                     /clone="Lm211"
                     /clone_lib "LmLV39cDNA"
                     /lab_host-"XL1-Blue MRF'"
     mRNA            <1..>250
BASE COUNT       56 a     78 c     73 g     43 t
ORIGIN

  Query Match             1.2%;  Score 22;  DB 27;  Length 250;
  Best Local Similarity  82.9%,  Pred. No. 1.34e-03;
  Matches    34,  Conservative    0,  Mismatches    6,  Indels    1;  Gaps    1;

Db     199 cctcaggtgaactgcccccaggtgcgtcacaacatgaagcg 238
           ||| |||||||||| ||||| | | |||||||  |||||||
Qy    1678 CCTGAGGTGAACTGCCCCACAGAG-GTCACAAGCTGAAGCG 1717


RESULT    10
ID   HS1164487  standard; RNA; EST; 326 BP.
AC   AA249798;
NI   g1881068
DT   13-MAR-1997 (Rel. 51, Created)
DT   13-MAR-1997 (Rel. 51, Last updated, Version 1)
DE   11371 seq F Human fetal heart, Lambda ZAP Express Homo sapiens
DE   cDNA 5'.
KW   EST.
OS   Homo sapiens (human)
OC   Eukaryotae; mitochondrial eukaryotes, Metazoa, Chordata;
OC   Vertebrata; Eutheria, Primates, Catarrhini; Hominidae; Homo.
RN   [1]
RP   1-326
RA   Liew C.C.;
RT   "cDNAs from human fetal heart (1997)";
RL   Unpublished.
CC   Contact  Liew CC Molecular Cardiology University of Toronto Banting
CC   Institute, 100 College St , Toronto, Ontario, M5G1L5 Tel:
CC   4169788758 Fax: 4169785650 Email. liewcc@utcc.utoronto.ca PCR
CC   PRimers FORWARD  5' GCCAAGCTCGAAATTAACCCTCACTAAAGGG 3' BACKWARD: 5'
CC   CCAGTGAATTGTAATACGACTCACTATAGGGCG 3' Seq primer: 5'
CC   GAAATTAACCCTCACTAAAGGG 3'.
FH   Key             Location/Qualifiers
```

```
FH
FT   source          1..326
FT                   /organism="Homo sapiens"
FT                   /note="Vector  Lambda ZAP Express, Site_1  EcoRI, Site_2
FT                   XhoI; mRNA was purified from human fetal hearts (8-10
FT                   weeks). cDNA was synthesized using a XhoI-Oligo dT
FT                   adaptor-primer. EcoRI adaptors were ligated, followed by
FT                   digestion with XhoI, fordirectional cloning into
FT                   predigested lambda ZAP Express."
FT                   /clone_lib "Human fetal heart, Lambda ZAP Express"
FT                   /lab_host="E. coli XL1-Blue"
FT   mRNA            <1..>326
SQ   Sequence 326 BP, 63 A, 87 C, 86 G, 90 T, 0 other,

  Query Match             1.2%;  Score 22,  DB 85,  Length 326,
  Best Local Similarity  78.9%;  Pred. No. 1.34e-03;
  Matches    30;  Conservative    0,  Mismatches   8,  Indels   0,  Gaps   0,

Db     257 cgtctgcagctcattatctgggcacaaggcggtaccca 294
           |||| |||  ||  | |||  ||||||||||||||||
Cp     785 CGTCGGCACTTCTCTGTCTCCGCACAAGGCGGTACCCA 748


RESULT    11
LOCUS       AA249798        326 bp    mRNA             EST       11-MAR-1997
DEFINITION  111371.seq.F Human fetal heart, Lambda ZAP Express Homo sapiens
            cDNA 5'.
ACCESSION   AA249798
NID         g1881068
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae, mitochondrial eukaryotes, Metazoa, Chordata,
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 326)
  AUTHORS   Liew,C.C.
  TITLE     cDNAs from human fetal heart (1997)
  JOURNAL   Unpublished (1997)
COMMENT
            Contact: Liew CC
            Molecular Cardiology
            University of Toronto
            Banting Institute, 100 College St., Toronto, Ontario, M5G1L5
            Tel: 4169788758
            Fax: 4169785650
            Email: liewcc@utcc.utoronto.ca
            PCR PRimers
            FORWARD: 5' GCCAAGCTCGAAATTAACCCTCACTAAAGGG 3'
            BACKWARD: 5' CCAGTGAATTGTAATACGACTCACTATAGGGCG 3'
            Seq primer: 5' GAAATTAACCCTCACTAAAGGG 3'.
FEATURES             Location/Qualifiers
     source          1..326
                     /organism "Homo sapiens"
                     /note="Vector  Lambda ZAP Express, Site_1  EcoRI, Site_2
                     XhoI; mRNA was purified from human fetal hearts (8-10
                     weeks).  cDNA was synthesized using a XhoI-Oligo dT
                     adaptor-primer.  EcoRI adaptors were ligated, followed by
                     digestion with XhoI, for directional cloning into
                     predigested lambda ZAP Express."
                     /clone_lib="Human fetal heart, Lambda ZAP Express"
                     /lab_host="E. coli XL1 Blue"
     mRNA            <1..>326
BASE COUNT       63 a     87 c     86 g     90 t
ORIGIN

  Query Match             1.2%;  Score 22,  DB 69,  Length 326,
  Best Local Similarity  78.9%;  Pred. No. 1.34e-03;
  Matches    30,  Conservative    0,  Mismatches   8,  Indels   0,  Gaps   0,

Db     257 cgtctgcagctcattatctgggcacaaggcggtaccca 294
           |||| |||  ||  | |||  ||||||||||||||||
Cp     785 CGTCGGCACTTCTCTGTCTCCGCACAAGGCGGTACCCA 748
```

```
RESULT    12
LOCUS       T67398          350 bp    mRNA             EST       23-JAN-1997
DEFINITION  lmESI0099 LmLV39cDNA Leishmania major cDNA clone lmi48 5' END.
ACCESSION   T67398
NID         g709782
KEYWORDS    EST.
SOURCE      Leishmania major.
  ORGANISM  Leishmania major
            Eukaryotae, mitochondrial eukaryotes, Euglenozoa, Kinetoplastida;
            Trypanosomatidae; Leishmania.
REFERENCE   1  (bases 1 to 350)
  AUTHORS   Levick,M.P., Blackwell,J.M., Connor,V., Coulson,R.M., Miles,A.,
            Smith,H.E., Wan,K.L. and Ajioka,J.W.
  TITLE     An expressed sequence tag analysis of a full length, spliced-leader
            cDNA library from Leishmania major promastigotes
  JOURNAL   Mol. Biochem. Parasitol. 76 (1-2), 345-348 (1996)
  MEDLINE   97077459
COMMENT
            Contact: Ajioka JW
            Laboratory for Parasite Genome Analysis
            Cambridge University
            Department of Pathology, Tennis Court Road, Cambridge CB2 1QP, UK.
            Tel: 01223333923
            Fax: 01223333923
            Email: jajioka@hgmp.mrc.ac.uk
            Seq primer: T3.
FEATURES             Location/Qualifiers
     source          1..350
                     /organism="Leishmania major"
                     /strain="LV39"
                     /note="Vector: Lambda UNIZAP; Site_1: NotI; Site_2: XhoI;
                     Mid or late log promastigotes Full length : splice leader
                     oligo used for second strand synthesis, directional
                     cloning NotI at 5' end, XhoI at 3' end."
                     /clone="lmi48"
                     /clone_lib="LmLV39cDNA"
                     /lab_host="XL1-Blue MRF'"
     mRNA            <1..>350
BASE COUNT       74 a    121 c     99 g     56 t
ORIGIN

  Query Match             1.2%;  Score 22;  DB 27;  Length 350;
  Best Local Similarity  82.9%;  Pred. No. 1.34e-03;
  Matches    34;  Conservative    0;  Mismatches   6;  Indels   1;  Gaps   1;

Db     209 cctgaggtgaactgccccaggtggtcacaagcatgaagca 248
           ||| |||||||||| ||||| | | |||||||  |||||||
Qy    1678 CCTGAGGTGAACTGGCCCACAGAG-GTCACAAGCTCAAGCG 1717


RESULT    13
LOCUS       AA234422        358 bp    mRNA             EST       03-MAR-1997
DEFINITION  zr70322.r1 Soares NbHMPu S1 Homo sapiens cDNA clone 668931 5'
            similar to gb:M14338 VITAMIN K-DEPENDENT PROTEIN S PRECURSOR
            (HUMAN);.
ACCESSION   AA234422
NID         g1858696
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae, mitochondrial eukaryotes, Metazoa, Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 358)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact. Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL , contact the
            IMAGE Consortium (info@image.llnl.gov) for further information
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 342.
FEATURES             Location/Qualifiers
     source          1..358
                     /organism="Homo sapiens"
                     /note="Organ: mixed (see below), Vector: pT7T3D-Pac
                     (Pharmacia) with a modified polylinker; Site_1: Not I,
                     Site_2: Eco RI. Equal amounts of plasmid DNA from three
                     normalized libraries (melanocyte 2NbHM, pregnant uterus
                     NbHPU, and fetal heart NbHH19W) were mixed, and ss circles
                     were made in vitro. Following HAP purification, this DNA
                     was used as tracer in a subtractive hybridization
                     reaction. The driver was PCR-amplified cDNAs from pools of
                     5,000 clones made from the same 3 libraries. The pools
                     consisted of I.M.A.G.E. clones 260232-265223,
                     340488-345479, and 484488-489479."
                     /clone="668931"
                     /clone_lib="Soares NbHMPu S1"
                     /tissue_type="Pooled human melanocyte, fetal heart, and
                     pregnant uterus"
                     /lab_host="DH10B"
     mRNA            <1..>358
BASE COUNT      106 a     57 c     83 g    112 t
ORIGIN

  Query Match               1.2%;  Score 23;  DB 65;  Length 358;
  Best Local Similarity  79.5%;  Pred. No. 7.42e-05;
  Matches   31;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Db       30 agctttggcttcatctttctcttgtgtggcttcacattt 68
            ||| | ||||||||||||||  |||||||    | |||||
Qy     1737 AGCCTGGGCTTCATCTTTCTCCTGTGTGGTCCCTCATTT 1775


RESULT   14
ID   HS1152547  standard; RNA; EST; 358 BP.
AC   AA234422;
NI   g1858696
DT   06-MAR-1997 (Rel. 51, Created)
DT   06-MAR-1997 (Rel. 51, Last updated, Version 1)
DE   zr72d02.r1 Soares NhHMPu S1 Homo sapiens cDNA clone 668931 5'
DE   similar to gb:M14338 VITAMIN K-DEPENDENT PROTEIN S PRECURSOR
DE   (HUMAN);.
KW   EST.
OS   Homo sapiens (human)
OC   Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
OC   Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   1-358
RA   Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,
RA   Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,
RA   Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,
RA   Trevaskis E., Waterston R., Williamson A., Wohldmann P., Wilson R.;
RT   "The WashU-Merck EST Project".
RL   Unpublished.
CC   Contact: Wilson RK WashU-Merck EST Project Washington University
CC   School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis,
CC   MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:
CC   est@watson.wustl.edu This clone is available royalty-free through
CC   LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for
CC   further information. Seq primer: -28m13 rev2 ET from Amersham High
CC   quality sequence stop: 342.
FH   Key             Location/Qualifiers
FH
FT   source          1..358
FT                   /organism="Homo sapiens"
FT                   /note="Organ: mixed (see below), Vector: pT7T3D-Pac
FT                   (Pharmacia) with a modified polylinker; Site_1: Not I;
FT                   Site_2: Eco RI. Equal amounts of plasmid DNA from three
FT                   normalized libraries (melanocyte 2NbHM, pregnant uterus
FT                   NbHPU, and fetal heart NbHH19W) were mixed, and ss circles
FT                   were made in vitro. Following HAP purification, this DNA
FT                   was used as tracer in a subtractive hybridization reaction
.
FT                   The driver was PCR-amplified cDNAs from pools of 5,000
FT                   clones made from the same 3 libraries. The pools consisted
FT                   of I.M.A.G.E. clones 260232-265223, 340488-345479, and
FT                   484488-489479."
FT                   /clone="668931"
FT                   /clone_lib="Soares NhHMPu S1"
FT                   /tissue_type="Pooled human melanocyte, fetal heart, and
FT                   pregnant uterus"
FT                   /lab_host="DH10B"
FT   mRNA            <1..>358
SQ   Sequence 358 BP; 106 A; 57 C; 83 G; 112 T; 0 other;

  Query Match               1.2%;  Score 23;  DB 83;  Length 358;
  Best Local Similarity  79.5%;  Pred. No. 7.42e-05;
  Matches   31;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Db       30 agctttggcttcatctttctcttgtgtggcttcacattt 68
            ||| | ||||||||||||||  |||||||    | |||||
Qy     1737 AGCCTGGGCTTCATCTTTCTCCTGTGTGGTCCCTCATTT 1775


RESULT   15
LOCUS       AA057786      382 bp    mRNA            EST       01-FEB-1997
DEFINITION  zf15e04.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone
            377022 3'.
ACCESSION   AA057786
NID         g1550437
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 382)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E.,
            Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1130   Std Error: 0.00
            Seq primer: -40M13 fwd. from Amersham
            High quality sequence stop: 272.
FEATURES             Location/Qualifiers
     source          1..382
                     /organism="Homo sapiens"
                     /note="Organ: heart; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCATCTTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library constructed by
                     M.Fatima Bonaldo. This library was constructed from the
                     same fetus as the fetal lung library, Soares fetal lung
                     NbHL19W."
                     /clone="377022"
                     /clone_lib="Soares fetal heart NbHH19W"
                     /sex="unknown"
                     /dev_stage="19 weeks"
                     /lab_host="DH10B (ampicillin resistant)"
     mRNA            complement(<1..>382)
BASE COUNT        70 a    103 c    127 g     79 t       3 others
ORIGIN

  Query Match             1.2%;  Score 22;  DB 30;  Length 382;
  Best Local Similarity  78.9%;  Pred. No. 1.34e-03;
  Matches    30;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Db    145 gctgctggaggcacagggccccgggctgggagtgctca 182
          |||||||||  ||  |  | ||||||  ||  |   ||||||||
Qy    850 GCTGCTGGACGCCCTGAGCCCCGAGCAGCTAGTGCTCA 887


Search completed: Thu Dec 18 21:17:38 1997
Job time : 533 secs.
```

```
*******************************************************************
                                                             (TM)
*******************************************************************

        Release 2.1D John F. Collins, Biocomputing Research Unit.
        Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
              Distribution rights by IntelliGenetics, Inc.

MPsrch_nn    n.a. - n.a. database search, using Smith-Waterman algorithm

Run on:           Thu Dec 18 19:49:10 1997   MasPar time 684.84 Seconds
                                             942.193 Million cell updates/sec
Tabular output not generated

Title:            >US-08-826-361A-20
Description:      (1-1257) from US08826361A.seq
Perfect Score:    1257
N.A. Sequence:        1 ATGAATTACAGCATTCCCAG ......GCCATCCGAGGTCTCCTGA 1257
     Comp                 TACTTAATGTCGTAAGGGTC       CGGTAGGCTCCAGAGGACT

Scoring table:    TABLE default
                  Gap 6

Nmatch   STD .    Dbase 0   Query 0

Searched:         707517 seqs  256659290 bases x 2

Post-processing   Minimum Match 0%
                  Listing first 45 summaries

Database:         EST-STS
                    1:EST1 2:EST2 3:EST3 4:EST4 5:EST5 6:EST6 7:EST7 8:EST8
                    9:EST9 10:EST10 11:EST11 12:EST12 13:EST13 14:EST14
                    15:EST15 16:EST16 17:EST17 18:EST18 19:EST19 20:EST20
                    21:EST21 22:EST22 23:EST23 24:EST24 25:EST25 26:EST26
                    27:EST27 28:EST28 29:EST29 30:EST30 31:EST31 32:EST32
                    33:EST33 34:EST34 35:EST35 36:EST36 37:EST37 38:EST38
                    39:EST39 40:EST40 41:EST41 42:EST42 43:EST43 44:EST44
                    45:EST45 46:EST46 47:EST47 48:EST48 49:EST49 50:EST50
                    51:EST51 52:EST52 53:EST53 54:EST54 55:EST55 56:EST56
                    57:EST57 58:EST58 59:EST59 60:EST60 61:EST61 62:EST62
                    63:EST63 64:EST64 65:EST65 66:EST66 67:EST67 68:EST68
                    69:EST69 70:EST70 71:EST71 72:EST72 73:EST73 74:EST74
                    75:EST75 76:EST76 77:EST77 78:EST78 79:EST79 80:EST80
                    81:EST81 82:EST82 83:EST83 84:EST84 85:EST85 86:EST86
                    87:EST87 88:EST88 89:EST89 90:EST90 91:EST91 92:EST92
                    93:EST93 94:EST94 95:EST95 96:EST96 97:EST97 98:EST98
                    99:EST99
Database:         EST-STS-TWO
                    100:EST100 101:EST101 102:EST102 103:EST103 104:EST104
                    105:EST105 106:EST106 107:EST107 108:EST108 109:EST109
                    110:EST110 111:EST111 112:EST112 113:EST113 114:EST114
                    115:EST115 116:EST116 117:EST117 118:EST118 119:EST119
                    120:EST120 121:EST121 122:EST122 123:EST123 124:EST124
                    125:EST125 126:EST126 127:EST127 128:EST128 129:EST129
                    130:EST130 131:EST131 132:EST132 133:EST133 134:EST134
                    135:EST135 136:EST136 137:EST137 138:EST138 139:EST139
                    140:EST140 141:EST141 142:EST142 143:EST143 144:EST144
                    145:EST145 146:EST146 147:EST147 148:EST148 149:EST149
                    150:EST150 151:EST151 152:EST152 153:EST153 154:EST154
                    155:EST155 156:EST156 157:EST157 158:EST158 159:EST159
                    160:EST160 161:EST161 162:EST162 163:EST163 164:EST164
                    165:EST165 166:EST166 167:EST167 168:EST168 169:EST169
                    170:EST170 171:EST171 172:EST172 173:EST173 174:EST174
                    175:EST175 176:EST176 177:EST177 178:EST178 179:EST179
                    180:EST180 181:EST181 182:EST182 183:EST183 184:EST184
                    185:EST185 186:EST186 187:EST187 188:EST188 189:EST189
                    190:EST190 191:EST191 192:EST192 193:EST193 194:EST194
                    195:EST195 196:EST196 197:EST197 198:EST198

Statistics:       Mean 11.184;  Variance 1.968;  scale 5.682

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                 SUMMARIES
                   %
Result           Query
   No.   Score   Match Length DB   ID           Description            Pred. No.
------------------------------------------------------------------------------
c    1      53     4.2    454  71  H82542       ys69h11.r1 Homo sapie  1.06e-56
c    2      52     4.1    390 149  W13715       ma83b04.r1 Soares mou  9.46e-55
c    3      52     4.1    4[illegible]  74  H91[illegible]  ys91d07.r1 Homo sapie  9.46e-55
     4      51     4.1    412 131  AA157597     zo44a11.r1 Stratagene  5.18e-53
     5      49     3.9    508 120  W53545       md56a06.r1 Soares mou  4.81e-47
     6      49     3.8    536 147  AA003987     mq80b08.r1 Soares mou  4.81e-47
     7      41     3.3    552 120  W53477       md45d07.r1 Soares mou  4.97e-34
     8      38     [illegible]    365 191  AA[illegible]2184  114[illegible] seq F Fetal he  1.12e-28
     9      38     [illegible]    [illegible]  [illegible]  [illegible]  yg[illegible].r1 Homo sapie  1.12e-28
    10      [illegible]     3.[illegible]    [illegible]4 195  AA[illegible]  [illegible].r1 Soares mou  1.12e-28
    11      [illegible]     2.9    56[illegible] 120  W53511       md45f07.r1 Soares mou  3.34e-25
    12      32     2.5    378 141  C08482       C.elegans cDNA clone   1.66e-18
    13      29     2.3    163   8  T83788       yd65a11.r1 Homo sapie  9.37e-14
    14      29     2.2    36[illegible]  89  CELK105GXF   C.elegans cDNA clone   3.13e-12
    15      29     2.2    421 137  AA044343     zk54a08.r1 Soares pre  3.13e-12
    16      29     2.2    [illegible] 1[illegible]  W74889       mo32d01.r1 Soares mou  3.13e-12
    17      27     2.1    368 198  N44459       SW3ICA1324SK Brugia m  9.67e-11
    18      27     2.1    48[illegible]   [illegible]  T[illegible]  yl11e[illegible].r1 Homo sapie  9.67e-11
    19      25     2.0    12[illegible] 141  C09766       C.elegans cDNA clone   7.15e-08
    20      25     2.0    2[illegible]  89  CELK082E1F   C.elegans cDNA clone   7.15e-08
    21      25     2.0    [illegible] 119  W39474       zc76d10.r1 Pancreatic  7.15e-08
    22      25     2.0    [illegible] 14[illegible]  T2[illegible]  NIBI014D12R Infant br  7.15e-08
    23      24     1.9    141  65  H2[illegible]  yl[illegible].r1 Homo sapie  1.69e-06
    24      24     1.9    [illegible] 154  AA011709     zi03d03.r1 Soares fet  1.69e-06
c   25      24     1.9    340 154  AA011660     zi03d03.s1 Soares fet  1.69e-06
    26      24     1.9    41[illegible]  4[illegible]  M7[illegible]  EST01[illegible] Homo sapiens  1.69e-06
    27      24     1.9    456 154  AA0[illegible]  zi03f11.r1 Soares fet  1.69e-06
    28      24     1.9    487  63  H23041       ym51d05.r1 Homo sapie  1.69e-06
    29      24     1.9    5[illegible] 151  W2[illegible]  mb96d08.r1 Soares mou  1.69e-06
    30      23     1.8    [illegible]  85  H5548[illegible]  [illegible] Homo sapien  3.58e-05
c   31      23     1.8    264  8[illegible]  T30[illegible]  EST1584[illegible] Homo sapiens  3.58e-05
    32      23     1.8    [illegible] 1[illegible]  HS2[illegible]  H. sapiens partial c[illegible]  3.58e-05
c   33      22     1.8    [illegible]  98  H66756       yr84b12.r1 Homo sapie  6.79e-04
    34      23     1.8    242  68  H74144       ys10b06.r1 Homo sapie  3.58e-05
c   35      23     1.8    344 158  AA034113     zi06d11.s1 Soares fet  3.58e-05
c   36      22     1.8    36[illegible]  35  RICR1514A    R.[illegible] cDNA, partial [illegible]  6.79e-04
    37      23     1.8    42[illegible]   [illegible]  T[illegible]  yl84[illegible].r1 Homo sapie  3.58e-05
c   38      23     1.8    431  15  T11624       A1324F Homo sapiens c  3.58e-05
    39      23     1.8    441  18  T51919       yb54a12.s1 Homo sapie  3.58e-05
    40      23     1.8    464 165  C17508       Human placenta cDNA 5  3.58e-05
c   41      22     1.8    408 1[illegible]  W49760       zc41h06.r1 Soares sen  6.79e-04
    42      22     1.8    47[illegible]  2[illegible]  R5[illegible]  yj59h12.s1 Homo sapie  6.79e-04
    43      23     1.8    485 173  W74443       zd75e11.s1 Soares fet  3.58e-05
c   44      23     1.8    573 116  W26710       13e1 Human retina cDN  3.58e-05
    45      23     1.8    594 154  AA0[illegible]  mg85e11.r1 Soares mou  3.58e-05


                                ALIGNMENTS

RESULT      1
LOCUS       H82542         454 bp    mRNA              EST       09-NOV-1995
DEFINITION  ys69h11.r1 Homo sapiens cDNA clone 220101 5' similar to
            gb:X51417_cds1 STEROID HORMONE RECEPTOR ERR2 (HUMAN);.
ACCESSION   H82542
NID         g1060631
KEYWORDS    EST.
```

```
SOURCE      human clone=220101 primer=M13RP1 library=Soares retina N2b4HR
            vector=pT7T3D (Pharmacia) with a modified polylinker host=DH10B
            (ampicillin resistant) Rsite1=Not I Rsite2=Eco RI 1st strand cDNA
            was primed with a Not I - oligo(dT) primer
            [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCGCTTTTTTTTTTTTTTTTTTT-3'],
            double-stranded cDNA was size selected, ligated to Eco RI adapters
            (Pharmacia), digested with Not I and cloned into the Not I and Eco
            RI sites of a modified pT7T3 vector (Pharmacia). The retinas were
            obtained from a 55 year old Caucasian male  and total cellular
            poly(A)+ RNA was extracted 6 hrs after their removal. The retina
            RNA was kindly provided by Roderick R. McInnes M.D. Ph.D. from the
            University of Toronto. Library constructed by Bento Soares and
            M.Fatima Bonaldo.
  ORGANISM  Homo sapiens
            Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
            Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
            Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
            Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 454)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops: 345
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
FEATURES             Location/Qualifiers
     source          1..454
                     /organism="Homo sapiens"
                     /clone="220101"
     mRNA            <1..>454
BASE COUNT       82 a    118 c    136 g    114 t       4 others
ORIGIN

  Query Match             4.2%;  Score 53;  DB 71;  Length 454;
  Best Local Similarity  65.7%;  Pred. No. 1.06e-56;
  Matches  119;  Conservative    0;  Mismatches  61;  Indels    1;  Gaps    1;

Db      137 cttcagcatcccacttgaggcatttcatgaagcgctggcaggcctgggagccttgcgcct 196
              ||||| ||| ||||||| |     || |||     |  |||||||||||||||  ||||||||
Cp      489 CTTCACCATTCCCACTTCGTAACACTTCCGAAGTCGGCAGGCCTGGCAGCTCTTGCGCCG 430

Db      197 ccatttggtgatctcgcactcgttggtggccgggcagctgtactcaatgttcccttggat 256
                 |||    |||     ||||  || || || || ||   | || ||| | |  ||||| ||
Cp      429 CTTTTTATCGATTGTACACTGATTTGTAGCTGGACAAATATAATCATTATGTCCTTGAAT 370

Db      257 agtcctcttgaagaaggccttgcaagcctcgcagagaccaacccctgagtggtagcccag 316
                | || || || |||||||| ||   | || || ||   |  || ||  |||| || || |
Cp      369 GCTTCTTTTAAAAAAGGCCTTACATCCTTCACACGAC-CAGACTCCATAGTGATATCCCG 311

Db      317 a 317
            |
Cp      310 A 310


RESULT      2
LOCUS       W13715         390 bp    mRNA             EST       10-SEP-1996
DEFINITION  ma83b04.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone 317263 5'
            similar to gb:L13740 EARLY RESPONSE PROTEIN NAK1 (HUMAN);.
ACCESSION   W13715
NID         g1297753
```

```
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 390)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: ETPrimer
            High quality sequence stop: 385.
FEATURES             Location/Qualifiers
     source          1..390
                     /organism="Mus musculus"
                     /note="Vector: pT7T3D (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCATTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library constructed by Bento
                     Soares and M.Fatima Bonaldo. RNA was kindly provided by
                     Dr. Minoru Ko (Wayne State University)."
                     /clone="317263"
                     /clone_lib="Soares mouse p3NMF19.5"
                     /dev_stage="19.5 dpc total fetus"
                     /lab_host="DH10B (ampicillin resistant)"
     mRNA            <1..>390
BASE COUNT       70 a    106 c    114 g    100 t
ORIGIN

  Query Match             4.1%;  Score 52;  DB 149;  Length 390;
  Best Local Similarity  62.7%;  Pred. No. 9.46e-55;
  Matches  128;  Conservative    0;  Mismatches  76;  Indels    0;  Gaps    0;

Db       80 cttaaccatcccgacactgagacacttctgaaatctgcagtactgacatcggtttcggcg 139
             ||| ||||| || ||     |   ||||||| ||| || ||||   ||| || |  || || ||
Cp      489 CTTCACCATTCCCACTTCGTAACACTTCCGAAGTCGGCAGGCCTGGCAGCTCTTGCGCCG 430

Db      140 tctcttgtccactggtcagttttatttgccaggcaaacatattttgcattttttctgcac 199
                | || || | ||  || |  ||  | ||  | |||| ||| |     || |    || |
Cp      429 GTTTTTATCGATTGTACACTGATTTGTAGCTGGACAAATATAATCATTATGTCCTTGAAT 370

Db      200 cgttctcttgaagaagccttgcagcctcgcagatgccgactccgtagtgctggacaggc 259
                |||| || || ||| |||| || || || || |   | |||||| ||||| |   |   |
Cp      369 GCTTCTTTTAAAAAAGGCCTTACATCCTTCACACGACCAGACTCCATAGTGATATCCCGA 310

Db      260 agcgttgtcgccgcacacggcaca 283
              ||||  |||| ||| || || ||
Cp      309 TGCGTAATCGCTGCAGACAGCGCA 286


RESULT      3
LOCUS       H91890         403 bp    mRNA             EST       29-NOV-1995
DEFINITION  ys81d07.r1 Homo sapiens cDNA clone 221197 5' similar to
            gb:X51417_cds1 STEROID HORMONE RECEPTOR ERR2 (HUMAN);.
```

```
ACCESSION   H91890
NID         g1087468
KEYWORDS    EST.
SOURCE      human clone=221197 primer=M13RP1 library=Soares retina N2b4HR
            vector=pT7T3D (Pharmacia) with a modified polylinker host=DH10B
            (ampicillin resistant) Rsite1=Not I Rsite2=Eco RI 1st strand cDNA
            was primed with a Not I - oligo(dT) primer
            [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCTTTTTTTTTTTTTTTTTTT-3'],
            double-stranded cDNA was size selected, ligated to Eco RI adapters
            (Pharmacia), digested with Not I and cloned into the Not I and Eco
            RI sites of a modified pT7T3 vector (Pharmacia). The retinas were
            obtained from a 55 year old Caucasian male  and total cellular
            poly(A)+ RNA was extracted 6 hrs after their removal. The retina
            RNA was kindly provided by Roderick R  McInnes M D  Ph D  from the
            University of Toronto. Library constructed by Bento Soares and
            M.Fatima Bonaldo.
  ORGANISM  Homo sapiens
            Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
            Deuterostomia, Chordata, Vertebrata; Gnathostomata; Osteichthyes;
            Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
            Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo
REFERENCE   1  (bases 1 to 403)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops: 330
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
FEATURES             Location/Qualifiers
     source          1..403
                     /organism="Homo sapiens"
                     /clone="221197"
     mRNA            <1..>403
BASE COUNT       73 a    113 c    118 g     97 t      2 others
ORIGIN

  Query Match             4.1%;  Score 52;  DB 74;  Length 403;
  Best Local Similarity  66.7%;  Pred. No. 9.46e-55;
  Matches  104;  Conservative    0;  Mismatches   52;  Indels    0;  Gaps    0;

Db     112 cttcagcatcccgactttgaggcatttcatgaagcgacaggctggcaggacttgcgcct 171
            ||||| ||| ||||||| |     || |||     |  ||||||||||||||  ||||||||
Qy     489 CTTCACTATTTTTATTTCGTAATATTTCCGAAGTCGGCAGGCTTGGCAGTTTTTGCCTCG 430

Db     172 ccgtttggtgatctcgcactcgttggtggccgggcagctgtactcaatgttcccttggat 231
               |||   |||    ||||  || || || || ||   | || ||| | |   |||| ||
Qy     429 GTTTTTATCGATTGTACATTGATTTGTAGTTGGATAAATATAATCATTATGTCCTTGAAT 370

Db     232 agtcctcttgaagaaggccttgcaagcctcgcagga 267
              | || || || |||||||| ||   | || || ||
Qy     369 GCTTCTTTTAAAAAAGGCCTTACATCCTTCACACGA 334


RESULT     4
LOCUS       AA157597      412 bp    mRNA            EST       11-DEC-1996
DEFINITION  zo44a02.r1 Stratagene endothelial cell 937223 Homo sapiens cDNA
            clone 589706 5' similar to gb X07282 RETINOIC ACID RECEPTOR BETA 2
            (HUMAN);.
ACCESSION   AA157597
NID         g1729205
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 412)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E.,
            Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 351.
FEATURES             Location/Qualifiers
     source          1..412
                     /organism="Homo sapiens"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI; Cloned unidirectionally.  Primer: Oligo dT.
                     Umbilical vein endothelial cells, passaged once. Average
                     insert size  1.0 kb; Uni-ZAP XR Vector;  5' adaptor
                     sequence  5' GAATTCGGCACGAG 3'  3' adaptor sequence: 5'
                     CTCGAGTTTTTTTTTTTTTTTTTTT 3'"
                     /clone="589706"
                     /clone_lib="Stratagene endothelial cell 937223"
                     /dev_stage="umbilical vein, 1 passage"
                     /lab_host="SOLR (kanamycin resistant)"
     mRNA            <1..>412
BASE COUNT      129 a     89 c    101 g     89 t      4 others
ORIGIN

  Query Match             4.1%;  Score 51;  DB 191;  Length 412;
  Best Local Similarity  64.9%;  Pred. No. 8.28e-53;
  Matches  111;  Conservative    0;  Mismatches   60;  Indels    0;  Gaps    0;

Db      29 gggtacccactatgggatcagcgcctgtgaaggatgtaagggctttttccgcagaagtatt 88
            || || |||||||| ||| |   | ||||| |||||||||| ||||||     ||||| ||
Qy     313 GGATATCACTATGGAGTCTGGTCGTGTGAAGGATGTAAGGCCTTTTTAAAAGAAGCATT 372

Db      89 cagaagaatatgatttacacttgtcaccgagataagaactgtgttattaatagaagtcacc 148
            ||     |||    ||| | |||||            ||  | |||    ||  ||||| |
Qy     373 TAAGGACATAATGATTATATTTGTCCAGCTACAAATCAGTGTACAATCGATAAAAACCGG 432

Db     149 aggaatcgatgccaatactgtcgactccagaagtgctttgaagtgggaatg 199
             | ||  | |||||   ||| ||||| | |||||| |  ||||||||||
Qy     433 CCCAAGAGTCCCCAGCCTCCCGATTTCGAAGTGTTACGAAGCCGGAACG 483


RESULT     5
LOCUS       W53545        508 bp    mRNA            EST       03-JUN-1996
DEFINITION  md56a06.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
            clone 372370 5' similar to gb:X57528 M.musculus mRNA for retinoic
            acid receptor-alpha (MOUSE);.
ACCESSION   W53545
NID         g1357435
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
            Murinae; Mus.
REFERENCE   1  (bases 1 to 508)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:233802
            Seq primer: mob.REGA+ET
            High quality sequence stop: 352.
FEATURES             Location/Qualifiers
     source          1..508
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGCAAATTTTTTTTTTTTTTTTTTTTT
                     T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
                     14.5dpc embryos [total RNA provided by  Minoru Ko, Wayne
                     State Univ., from 2 ]; double-stranded cDNA was ligated to
                     Eco RI adaptors (Pharmacia), digested with Not I and
                     cloned into the Not I and Eco RI sites of the modified
                     pT7T3 vector. Library went through one round of
                     normalization, and was constructed by Bento Soares and
                     M.Fatima Bonaldo."
                     /clone="372370"
                     /clone_lib="Soares mouse embryo NbME13.5 14.5"
                     /sex="unknown"
                     /tissue_type="embryo"
                     /dev_stage="13.5-14.5dpc total fetus"
                     /lab_host="DH10B"
     mRNA            <1..>508
BASE COUNT      123 a    168 c    127 g     90 t
ORIGIN

  Query Match             3.8%;  Score 48;  DB 120;  Length 508;
  Best Local Similarity  63.6%;  Pred. No. 4.81e-47;
  Matches  112;  Conservative    0;  Mismatches  64;  Indels    0;  Gaps    0;

Db     233 catccggctaccactatggagtcagcgcctgtgaaggctgtaaggcttcttccgacgaa 292
           |||| || || |||||||| ||| |  | ||||| || ||||||| ||| ||    | |||
Qy     308 CATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATGTAAGGCCTTTTTTAAAAGAA 367

Db     293 gcatccagaagaacatggtgtatacgtgtcaccgggacaagaactgcatcatcaacaagg 352
           |||| ||       | |  | |||| ||||            ||  | || |  ||| | ||
Qy     368 GCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCAGTGTACAATCGATAAAA 427

Db     353 tgacccggaaccgctgccagtactgccggctgcagaaatgtttcgacgtgggcatg 408
                || ||  ||||||||  |||||| || | ||| |||| ||| ||||| |||
Qy     428 ACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTACGAAGTGGGAATG 483


RESULT      6
LOCUS       AA003987       536 bp    mRNA            EST       23-JUL-1996
DEFINITION  mg80b08.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
            clone 439287 5' similar to gb:X57528 M.musculus mRNA for retinoic
            acid receptor-alpha (MOUSE);.
ACCESSION   AA003987
NID         g1447462
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
            Murinae; Mus.
REFERENCE   1  (bases 1 to 536)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:264623
            Seq primer: mob.REGA+ET
            High quality sequence stop: 346.
FEATURES             Location/Qualifiers
     source          1..536
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGGAAATTTTTTTTTTTTTTTTTTTTT
                     T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
                     14.5dpc embryos [total RNA provided by  Minoru Ko, Wayne
                     State Univ., from 2 ]; double-stranded cDNA was ligated to
                     Eco RI adaptors (Pharmacia), digested with Not I and
                     cloned into the Not I and Eco RI sites of the modified
                     pT7T3 vector. Library went through one round of
                     normalization, and was constructed by Bento Soares and
                     M.Fatima Bonaldo."
                     /clone="439287"
                     /clone_lib="Soares mouse embryo NbME13.5 14.5"
                     /sex="unknown"
                     /tissue_type="embryo"
                     /dev_stage="13.5-14.5dpc total fetus"
                     /lab_host="DH10B"
     mRNA            <1..>536
BASE COUNT      129 a    180 c    131 g     96 t
ORIGIN

  Query Match             3.8%;  Score 48;  DB 147;  Length 536;
  Best Local Similarity  63.6%;  Pred. No. 4.81e-47;
  Matches  112;  Conservative    0;  Mismatches  64;  Indels    0;  Gaps    0;

Db     286 catccggctaccactatggagtcagcgcctgtgaaggctgtaaggcttcttccgacgaa 345
           |||| || || |||||||| ||| |  | ||||| || ||||||| ||| ||    | |||
Qy     308 CATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATGTAAGGCCTTTTTTAAAAGAA 367

Db     346 gcatccagaagaacatggtgtatacgtgtcaccgggacaagaactgcatcatcaacaagg 405
           |||| ||       | |  | |||| ||||            ||  | || |  ||| | ||
Qy     368 GCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCAGTGTACAATCGATAAAA 427

Db     406 tgacccggaaccgctgccagtactgccggctgcagaaatgtttcgacgtgggcatg 461
                || ||  ||||||||  |||||| || | ||| |||| ||| ||||| |||
Qy     428 ACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTACGAAGTGGGAATG 483


RESULT      7
LOCUS       W53477         552 bp    mRNA            EST       03-JUN-1996
DEFINITION  md45d07.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
            clone 371341 5' similar to gb:X07282 RETINOIC ACID RECEPTOR BETA-2
            (HUMAN); gb:X57528 M.musculus mRNA for retinoic acid receptor-alpha
            (MOUSE);.
ACCESSION   W53477
```

```
NID          g1357321
KEYWORDS     EST.
SOURCE       house mouse.
  ORGANISM   Mus musculus
             Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
             Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
             Murinae; Mus.
REFERENCE    1  (bases 1 to 552)
  AUTHORS    Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
             Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
             Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
             Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
             Waterston,R.
  TITLE      The WashU-HHMI Mouse EST Project
  JOURNAL    Unpublished (1996)
COMMENT
             Contact: Marra M/Mouse EST Project
             WashU-HHMI Mouse EST Project
             Washington University School of MedicineP
             4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
             Tel: 314 286 1800
             Fax: 314 286 1810
             Email: mouseest@watson.wustl.edu
             This clone is available royalty-free through LLNL ; contact the
             IMAGE Consortium (info@image.llnl.gov) for further information.
             MGI:232773
             Seq primer: mob.REGA+ET
             High quality sequence stop: 319.
FEATURES             Location/Qualifiers
     source          1..552
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI. 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGGAAATTTTTTTTTTTTTTTTTTTTTTT
                     T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
                     14.5dpc embryos [total RNA provided by  Minoru Ko, Wayne
                     State Univ., from 2 ]; double-stranded cDNA was ligated to
                     Eco RI adaptors (Pharmacia), digested with Not I and
                     cloned into the Not I and Eco RI sites of the modified
                     pT7T3 vector. Library went through one round of
                     normalization, and was constructed by Bento Soares and
                     M.Fatima Bonaldo."
                     /clone="371341"
                     /clone_lib="Soares mouse embryo NbME13.5 14.5"
                     /sex="unknown"
                     /tissue_type="embryo"
                     /dev_stage="13.5-14.5dpc total fetus"
                     /lab_host="DH10B"
     mRNA            <1..>552
BASE COUNT      140 a    151 c    168 g     93 t
ORIGIN

  Query Match             3.3%;  Score 41;  DB 120;  Length 552;
  Best Local Similarity  62.2%;  Pred. No. 4.97e-34;
  Matches  130;  Conservative    0;  Mismatches  77;  Indels    2;  Gaps    2;

Db     230 ctacttcgtttaccaggacaagtcatcaagctaccactatgacgttagtgcctgccgagg 289
           ||||    || |||    ||   |   ||||||| ||  |||||||| ||| |   |  ||  || ||
Qy     285 CTGCCCTGTCTGCAGCGATTACGTATCGGGATATCATTATGGAGCCTGGTCCGTGGAAGG 344

Db     290 gtgcaagggctttttccgcagaagtattca-gaagaacatgatctacacttgccatcgag 348
            || |||| ||.|||     ||||| ||||| | | |  |||||        |  ||| | |
Qy     345 ATGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTAC 404

Db     349 ataagaactgcgtcattaacaaggtcactaggaacgatgccagtactgacgacctgcaga 408
           | ||   | ||     ||   | ||     |     | ||  | ||||||  |||||| || | ||
Qy     405 A-AATCAGTGTACAATCGATAAAAATCGGCGTAAGATTCGCCAGGCTTGCCGACTTCGGA 463

Db     409 agtgctttgaagtgggcatgtcaaagagt 437
           |||| |   ||'|| || |||    ||| ||
Qy     464 AGTGTTACGAAGTGGGAATGGTGAAGTGT 492
```

```
RESULT     8
LOCUS       AA092184      365 bp    mRNA            EST       24-OCT-1996
DEFINITION  114699 seq F Fetal heart, Lambda ZAP Express Homo sapiens cDNA 5'.
ACCESSION   AA092184
NID         g1637173
KEYWORDS    EST.
SOURCE      human
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 365)
  AUTHORS   Liew,C.C.
  TITLE     cDNAs from fetal heart (1996)
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Liew CC
            Molecular Cardiology
            University of Toronto
            Banting Institute, 100 College St , Toronto, Ontario, M5G1L5
            Tel: 4169788758
            Fax: 4169785650
            Email: liewcc@utcc.utoronto.ca
            PCR PRimers
            FORWARD: 5' GCCAAGCTCGAAATTAACCCTCACTAAAGGG 3'
            BACKWARD: 5' CCAGTGAATTGTAATACGACTCACTATAGGGCG 3'
            Seq primer: 5' GAAATTAACCCTCACTAAAGGG 3'.
FEATURES             Location/Qualifiers
     source          1..365
                     /organism="Homo sapiens"
                     /note="Vector: Lambda ZAP Express; Site_1: EcoRI; Site_2:
                     XhoI, mRNA was purified from human fetal hearts (8-10
                     weeks).  cDNA was synthesized using a XhoI-Oligo dT
                     adaptor-primer.  EcoRI adaptors were ligated, followed by
                     digestion with XhoI, for directional cloning into
                     predigested lambda ZAP Express."
                     /clone_lib="Fetal heart, Lambda ZAP Express"
                     /lab_host="E. coli XL1-Blue"
     mRNA            <1..>365
BASE COUNT       88 a     95 c    103 g     89 t
ORIGIN

  Query Match             3.0%;  Score 38;  DB 181;  Length 365;
  Best Local Similarity  77.1%;  Pred. No. 1.12e-28;
  Matches   54;  Conservative    0;  Mismatches  16;  Indels    0;  Gaps    0;

Db     164 cccttcagtgaagcttcgatgatgggcttactgaccaacctggcagacagggagctggtt 223
           |||||||  || || || |||||    |  ||||||||  |||| |||| |||| ||||
Qy     703 CCCTTCACCGAGGCCTCCATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTA 762

Db     224 cacatgatca 233
           ||||||||||
Qy     763 CACATGATCA 772


RESULT     9
LOCUS       R52304        393 bp    mRNA            EST       18-MAY-1995
DEFINITION  yg79c04.r1 Homo sapiens cDNA clone 39602 5' similar to gb:M29960
            STEROID RECEPTOR TR2 (HUMAN);.
ACCESSION   R52304
NID         g814206
KEYWORDS    EST.
SOURCE      human clone=39602 library=Soares infant brain 1NIB vector=Lafmid BA
            host=DH10B (ampicillin resistant) primer=M13RP1 Rsite1=Not I
            Rsite2=Hind III Whole brain from a 73 days post natal female. 1st
            strand cDNA was primed with a Not I - oligo(dT) primer [5'
            AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTTTT 3']; double-stranded
            cDNA was ligated to Hind III adaptors (Pharmacia), digested with
            Not I and directionally cloned into the Not I and Hind III sites of
            the Lafmid BA vector. Library went through one round of
            normalization. Library constructed by Bento Soares and M.Fatima
            Bonaldo.
  ORGANISM  Homo sapiens
            Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
            Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
            Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
            Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo
REFERENCE   1  (bases 1 to 393)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            GDB: G00-412 143
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops: 303
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
FEATURES             Location/Qualifiers
     source          1..393
                     /organism="Homo sapiens"
                     /clone="39602"
BASE COUNT      122 a     88 c     84 g     97 t      2 others
ORIGIN

  Query Match              3.0%;  Score 38;  DB 26;  Length 393;
  Best Local Similarity  66.7%;  Pred. No. 1.12e-29;
  Matches    74;  Conservative    0;  Mismatches   37;  Indels    0;  Gaps    0;

Db     271 aggaccaaataaggttttgatctttgcgtagtatgtggagacaaagcatcaggacgtca 330
           |||  |||| | || |    | |   | ||||   || ||   | ||  | ||||| |||   |||
Qy     261 AGGTTCAAAGAGGGATGCTCACTTCTGCGCTGTCTGCAGCGATTACGCATCGGGATATCA 320

Db     331 ttatggagcagtaacttgtgaaggctgcaaaggtttttttaaaagnagcat 381
             |||||||        | ||||||||| || || |   |||||||||| |||||
Qy     321 CTATGGAGTCTGGTCGTGTGAAGGATGTAAGGCCTTTTTTAAAAGAAGCAT 371


RESULT    10
LOCUS       AA049907        604 bp    mRNA             EST       30-DEC-1996
DEFINITION  mj13e05.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
            clone 475996 5' similar to gb:M73779 RETINOIC ACID RECEPTOR ALPHA-1
            (HUMAN); gb:X57528 M.musculus mRNA for retinoic acid receptor-alpha
            (MOUSE);.
ACCESSION   AA049907
NID         g1529596
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 604)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, Mo 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:286740
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 359.
FEATURES             Location/Qualifiers
     source          1..604
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker. Site_1: Not I. Site_2: Eco RI. 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGGAAATTTTTTTTTTTTTTTTTTTTTT
                     T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
                     14.5dpc embryos [total RNA provided by  Minoru Ko, Wayne
                     State Univ., from 2 ]; double-stranded cDNA was ligated to
                     Eco RI adaptors (Pharmacia), digested with Not I and
                     cloned into the Not I and Eco RI sites of the modified
                     pT7T3 vector. Library went through one round of
                     normalization, and was constructed by Bento Soares and
                     M.Fatima Bonaldo."
                     /clone="475996"
                     /clone_lib="Soares mouse embryo NbME13.5 14.5"
                     /sex="unknown"
                     /tissue_type="embryo"
                     /dev_stage="13.5-14.5dpc total fetus"
                     /lab_host="DH10B"
     mRNA            <1..>604
BASE COUNT      142 a    194 c    153 g    115 t
ORIGIN

  Query Match              3.0%;  Score 38;  DB 195;  Length 604;
  Best Local Similarity  60.8%;  Pred. No. 1.12e-28;
  Matches   107;  Conservative    0;  Mismatches   69;  Indels    0;  Gaps    0;

Db     287 catccggctaccactatggggtcagcgcctatgaggctgtaagggcttcttccgacgaa 346
           |||| || || |||||||| ||| |   | ||||| || ||||||| ||| ||     | | |
Qy     308 CATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATGTAAGGCCTTTTTTAAAAGAA 367

Db     347 gcatccagaagaacatggtgtatacgtgtcacccggacaagactgcatcatcaacaagg 406
           |||| ||      | |  |  ||||  ||||            ||  | || |   ||| | ||
Qy     368 GCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCAGTGTACAATCGATAAAA 427

Db     407 tgacccggaacgggctgcagtactgccggctgcaggaatgtttcgacgttgggatg 462
                 || ||  |    |||  |||||| || | ||| |||| ||| || || |||
Qy     428 ACCGGGGAAGAAGCTGGCAGGCCTGCCGACTTCGGAAGTGTTACCAAGTCGGAATG 483


RESULT    11
LOCUS       W53511          566 bp    mRNA             EST       03-JUN-1996
DEFINITION  md45f07.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
            clone 371365 5' similar to gb:X07282 RETINOIC ACID RECEPTOR BETA-2
            (HUMAN); gb:X57528 M.musculus mRNA for retinoic acid receptor alpha
            (MOUSE);.
ACCESSION   W53511
NID         g1357336
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
            Murinae; Mus.
REFERENCE   1  (bases 1 to 566)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
```

```
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway. Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information
            MGI:232797
            Seq primer: mob.REGA+ET
            High quality sequence stop: 360.
FEATURES             Location/Qualifiers
     source          1..566
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker, Site_1. Not I, Site_2. Eco RI, 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGGAAATTTTTTTTTTTTTTTTTTTTTT
                     T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
                     14.5dpc embryos [total RNA provided by  Minoru Ko, Wayne
                     State Univ.  from 2 ]; double-stranded cDNA was ligated to
                     Eco RI adaptors (Pharmacia), digested with Not I and
                     cloned into the Not I and Eco RI sites of the modified
                     pT7T3 vector. Library went through one round of
                     normalization, and was constructed by Bento Soares and
                     M.Fatima Bonaldo."
                     /clone="371365"
                     /clone_lib="Soares mouse embryo NbME13.5 14.5"
                     /sex="unknown"
                     /tissue_type="embryo"
                     /dev_stage="13.5-14.5dpc total fetus"
                     /lab_host="DH10B"
     mRNA            <1..>566
BASE COUNT      142 a    158 c    170 g     96 t
ORIGIN

  Query Match             2.9%;  Score 36;  DB 120;  Length 566;
  Best Local Similarity  62.7%;  Pred. No. 3.34e-25;
  Matches  126;  Conservative    0;  Mismatches  72;  Indels    3;  Gaps    3;

Db      242 [illegible] 301
            ||||   || |||   ||  |  ||||||| || |||||||| ||| |   | || || ||
Qy      285 CTGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGC 344

Db      302 gtgcaaggcttttttccgcagaagtattca-gaagaacatgatctacacttgccatcgag 360
             || |||| ||||||     ||||| |||||  | | |   |||||       |  ||| | |
Qy      345 ATGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTAC 404

Db      361 atcagactgcgtcattaacaaggtcactaggaaccgatgccagtactgcccgcttgcag 420
            | ||  | ||    ||  | ||    |     | ||  | ||||||  |||||| ||| | |
Qy      405 A-AATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTT-CGG 462

Db      421 aagtgctttgaagtgggcatg 441
            ||||| |  |||||||| |||
Qy      463 AAGTGTTACGAACTGGGAATG 483


RESULT   12
LOCUS       C08482          378 bp    mRNA             EST       09 SEP 1996
DEFINITION  C.elegans cDNA clone yk77h11 . 5' end, single read.
ACCESSION   C08482
NID         g1533553
KEYWORDS    EST(expressed sequence tag).
SOURCE      Caenorhabditis elegans (strain N2) embryo hermaphrodite cDNA to
            mRNA, clone_lib:Yuji Kohara unpublished cDNA clone:yk77h11
  ORGANISM  Caenorhabditis elegans
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Nematoda;
            Secernentea; Rhabditia; Rhabditida; Rhabditina; Rhabditoidea;
            Rhabditidae; Caenorhabditis.
REFERENCE   1  (bases 1 to 378)
  AUTHORS   Kohara,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-APR-1996) to the DDBJ/EMBL/GenBank databases. Yuji
            Kohara, National Institute of Genetics, Gene Library Lab; Yata
            1111, Mishima, Shizuoka 411, Japan (E-mail:ykohara@lab.nig.ac.jp,
            Tel:0559-81-6854, Fax:0559-81-6855)
REFERENCE   2  (sites)
  AUTHORS   Kohara,Y., Motohashi,T., Tabara,H., Watanabe,H., Sugimoto,A.,
            Sano,M., Miyata,A. and Nishigaki,A.
  TITLE     Expression map of the C.elegans genome
  JOURNAL   Unpublished (1996)
FEATURES             Location/Qualifiers
     source          1..378
                     /organism="Caenorhabditis elegans"
                     /strain="N2"
                     /clone="yk77h11"
                     /clone_lib="Yuji Kohara unpublished cDNA"
                     /dev_stage="embryo"
                     /sex="hermaphrodite"
BASE COUNT      126 a     75 c     89 g     82 t      6 others
ORIGIN

  Query Match             2.5%;  Score 32;  DB 141;  Length 378;
  Best Local Similarity  65.2%;  Pred. No. 1.66e-18;
  Matches   58;  Conservative    0;  Mismatches  31;  Indels    0;  Gaps    0;

Db      122 ttttgngcagtttgnggggatacggcactgggaaagcactacggtgtaacggcttgtnat 181
            || || || || ||   | |||    |||  |||| | ||||| || ||    | | |||
Qy      283 TTCTGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAA 342

Db      182 ggatgcaagggatttnncagaagaagtat 210
            ||||| ||||  |||    | |||||| ||
Qy      343 GGATGTAAGGCCTTTTTTAAAAGAAGCAT 371


RESULT   13
LOCUS       T83788          163 bp    mRNA             EST       16-MAR-1995
DEFINITION  yd65a11.r1 Homo sapiens cDNA clone 113084 5' similar to gb:M73779
            RETINOIC ACID RECEPTOR ALPHA-1 (HUMAN);.
ACCESSION   T83788
NID         g712076
KEYWORDS    EST.
SOURCE      human clone=113084 library=Soares fetal liver spleen 1NFLS
            vector=pT7T3 (Pharmacia) with a modified polylinker host=DH10B
            (ampicillin resistant) primer=M13RP1 Rsite1=Pac I Rsite2=Eco RI
            Liver and spleen from a 20 week-post conception male fetus. 1st
            strand cDNA was primed with a Pac I - oligo(dT) primer [5'
            AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTT 3'], double-stranded
            cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac
            I and cloned into the Pac I and Eco RI sites of the modified pT7T3
            vector.  Library went through one round of normalization. Library
            constructed by Bento Soares and M.Fatima Bonaldo.
  ORGANISM  Homo sapiens
            Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 163)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence starts: 1
            High quality sequence stops: 1
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality.
FEATURES             Location/Qualifiers
     source          1..163
                     /organism="Homo sapiens"
                     /clone="113084"
BASE COUNT       50 a     41 c     49 g     23 t
ORIGIN

  Query Match             2.3%;  Score 29;  DB 8;  Length 163;
  Best Local Similarity  78.4%;  Pred. No. 9.37e-14;
  Matches   40;  Conservative    0;  Mismatches  11;  Indels    0;  Gaps    0;

Db       35 cggaaccgctgccagtactgccgactgcagaagtgctttgaagtgggcatg 85
            || ||  ||||||||  ||||||||| | |||||| |   |||||||| |||
Qy      433 CGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTACGAAGTGGGAATG 483


RESULT    14
LOCUS       CELK105GXF     360 bp    mRNA            EST       13-DEC-1995
DEFINITION  C.elegans cDNA clone yk105g10 : 5' end, single read.
ACCESSION   D75525
NID         g1121318
KEYWORDS    EST(expressed sequence tag)
SOURCE      Caenorhabditis elegans (strain N2, ) (library: Yuji kohara
            unpublished cDNA) Hermaphrodite embryo embryo cDNA to mRNA
  ORGANISM  Caenorhabditis elegans
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Nematoda;
            Secernentea, Rhabditia, Rhabditida; Rhabditina; Rhabditoidea;
            Rhabditidae; Caenorhabditis
REFERENCE   1  (bases 1 to 360)
  AUTHORS   Kohara,Y., Motohashi,T., Tabara,H., Sugimoto,A., Watanabe,H. and
            Nishigaki,A
  TITLE     Toward an expression map of the C.elegans genome
  JOURNAL   Unpublished (1995)
COMMENT     Submitted (23-Aug-1995) to DDBJ by.
            Yuji Kohara
            Gene Library Lab.
            National Institute of Genetics
            Yata 1111, Mishima Shizuoka
            411 Japan
            Phone: 0559-81-6854
            Fax  : 0559-81-6855
            Email:ykohara@ddbj.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1..360
                     /organism="Caenorhabditis elegans"
                     /strain="N2"
                     /dev_stage="embryo"
                     /sex="Hermaphrodite"
                     /tissue_type="embryo"
                     /clone_lib="Yuji Kohara unpublished cDNA"
BASE COUNT       95 a     86 c    104 g     65 t     10 others
ORIGIN

  Query Match             2.2%;  Score 28;  DB 89;  Length 360;
  Best Local Similarity  65.5%;  Pred. No. 3.13e-12;
  Matches   57;  Conservative    0;  Mismatches  30;  Indels    0;  Gaps    0;

Db       98 tgcgctgtntgcggggacgtggctctcggaaaacactacggagtgaacgcttgcaacggg 157
            |||||||| ||| | ||     ||     ||| | ||||| |||||      | ||  | ||
Qy      286 TGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGA 345

Db      158 tgcaaaggattttttaggagaagcatt 184
            || || |  |||||||  |||||||||
Qy      346 TGTAAGGCCTTTTTTAAAAGAAGCATT 372


RESULT    15
LOCUS       AA044343       431 bp    mRNA            EST       04-SEP-1996
DEFINITION  zk54a08.r1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone
```

```
            486614 5'.
ACCESSION   AA044343
NID         g1522199
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 431)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 358.
FEATURES             Location/Qualifiers
     source          1..431
                     /organism="Homo sapiens"
                     /note="Organ: uterus; Vector: pT7T3-Pac; Site_1: Not I;
                     Site_2: Eco RI; 1st strand cDNA was primed with a Not I -
                     oligo(dT) primer [5'
                     AACTGGAAGAATTCGCGGCCGCTTTTTTTTTTTTTTTTTTT 3'],
                     double stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization. Library
                     constructed by M. Fatima Bonaldo."
                     /clone="486614"
                     /clone_lib="Soares pregnant uterus NbHPU"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="DH10B"
     mRNA            <1..>431
BASE COUNT       75 a    136 c    143 g     71 t      6 others
ORIGIN

  Query Match             2.2%;  Score 28;  DB 137;  Length 431;
  Best Local Similarity  75.9%;  Pred. No. 3.13e-12;
  Matches   41;  Conservative    0;  Mismatches  13;  Indels    0;  Gaps    0;

Db      192 [illegible] 245
            ||| |||||||||||||||  ||||| |||| |   | |   | | ||||||| ||
Qy      621 GGACGCCCTGAGCCCCGAGCAGCTAGTGCTCACCCTCCTGGAGGCTGAGCCGCC 674


Search completed: Thu Dec 18 20:00:47 1997
Job time : 697 secs.
```

```
******************************************************************

                        MPSRCH                          (TM)

******************************************************************

        Release 2.1D John F. Collins, Biocomputing Research Unit
        Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
              Distribution rights by IntelliGenetics, Inc.

MPsrch_nn  nucleic - nucleic database search, using Smith-Waterman algorithm

Run on:           Thu Dec 18 20:01:10 1997; MasPar time 355.48 Seconds
                                            971.750 Million cell updates/sec
Tabular output not generated

Title:            >US-08-826-361A-20
Description:      (1-1257) from US08826361A.seq
Perfect Score:    1257
N.A. Sequence       1 ATGAATTACAGCATTCCCAG         GGCATGGGAGGTCTGGCTGA 1257
    Comp:             TACTTAATGTCGTAAGGGTC         CCGTACCCTCCAGACCGACT

Scoring table:    TABLE default
                  Gap 6

Nmatch   STD      Phase 0   Query 0

Searched:         359985 seqs, 137405154 bases x 2

Post-processing:  Minimum Match 0%
                  Listing first 45 summaries

Database:         EST-STS-THREE
                    1:EST199 2:EST200 3:EST201 4:EST202 5:EST203 6:EST204
                    7:EST205 8:EST206 9:EST207 10:EST208 11:EST209 12:EST210
                    13:EST211 14:EST212 15:EST213 16:EST214 17:EST215
                    18:EST216 19:EST217 20:EST218 21:EST219 22:EST220
                    23:EST221 24:EST222 25:EST223 26:EST224 27:EST225
                    28:EST226 29:EST227 30:EST228 31:EST229 32:EST230
                    33:EST231 34:EST232 35:EST233 36:STS1 37:STS2 38:STS3
                    39:STS4 40:STS5 41:STS6 42:STS7 43:STS8 44:STS9 45:STS10
                    46:STS11 47:STS12 48:STS13
Database:         EST-STS-FOUR
                    49:gnEST1 50:gnEST2 51:gnEST3 52:gnEST4 53:gnEST5
                    54:gnEST6 55:gnEST7 56:gnEST8 57:gnEST9 58:gnEST10
                    59:gnEST11 60:gnEST12 61:gnEST13 62:gnEST14 63:gnEST15
                    64:gnEST16 65:gnEST17 66:gnEST18 67:gnEST19 68:gnEST20
                    69:gnEST21 70:gnEST22 71:gnEST23 72:gnEST24 73:gnEST25
                    74:gnSTS 75:enEST1 76:enEST2 77:enEST3 78:enEST4
                    79:enEST5 80:enEST6 81:enEST7 82:enEST8 83:enEST9
                    84:enEST10 85:enEST11 86:enEST12 87:enEST13 88:enEST14
                    89:enEST15 90:enEST16 91:enEST17 92:enEST18 93:enEST19
                    94:enEST20 95:enEST21 96:enEST22 97:enEST23 98:enEST24
                    99:enEST25 100:enEST26 101:enEST27 102:enEST28
                    103:enEST29 104:enEST30 105:enEST31 106:enSTS 107:ueEST1
                    108:ueEST2

Statistics:       Mean 11.210; Variance 2.020; scale 5.549

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES
                   %
Result           Query
```

```
     No.  Score  Match Length DB  ID            Description           Pred. No.
   -----------------------------------------------------------------------------
      1     53    4.2   2202 44  G28564        human STS SHGC-35295.  1.70e-55
 c    2     52    4.1    [illegible]  MM7156        ma27b04.r1 Soares mou  1.41e-[illegible]
 c    3     41    3.3    5[illegible] 1[illegible]  AA[illegible]        mh50b06.r1 Soares mou  1.35e-31
      4     26    2.1    [illegible] 2[illegible]  N59615        yv57g05.r1 Soares fet  4.29e-09
      5     24    1.9    156 36  DM36D7S       D. melanogaster STS d  1.72e-06
      6     24    1.9    536 97  MM94013       mb96d08.r1 Soares mou  1.71e-06
 c    7     22    1.8    326 85  HS1164487     111371.seq.F Human fe  5.95e-04
 c    8     22    1.8    326 60  AA249798      111371.seq.F Human fe  5.95e-04
      9     22    1.8    392 11  AA[illegible]     zf15e04.s1 Soares fet  5.95e-04
 c   10     22    1.8    399 21  AA021604      ze67e04.s1 Soares ret  5.95e-04
     11     22    1.8    424 89  HSW0986       ze08h09.r1 Soares fet  5.95e-04
     12     22    1.8    424  2  W86[illegible]        ze08h09.r1 Soares fet  5.95e-04
     13     22    1.8    452 18  AA[illegible]       ze76d08.s1 Soares fet  5.95e-04
 c   14     23    1.8    490 16  AA190298      zp95e10.s1 Stratagene  3.38e-05
 c   15     22    1.8    498 11  AA179191      zp46b09.r1 Stratagene  5.95e-04
 c   16     22    1.8    501 83  HS1151180     zr69e09.r1 Soares NhH  5.95e-04
 c   17     22    1.8    501 65  AA233250      zr69e09.r1 Soares NhH  5.95e-04
 c   18     22    1.8    540 13  AA[illegible]284      LD02433.5prime LD Dro  5.95e-04
     19     21    1.7    295 26  SSZ84056      S scrofa mRNA, expres  9.28e-03
 c   20     21    1.7    314 44  G27335        human STS CHLC.ATA70B  9.28e-03
     21     21    1.7    320  2  W95500        ze02b04.r1 Soares fet  9.28e-03
     22     21    1.7    320 90  HSW5005       ze02b04.r1 Soares fet  9.28e-03
 c   23     21    1.7    359 36  G00807        fruit fly STS Dm0760   9.28e-03
     24     21    1.7    384 66  AA235237      zs38g05.r1 Soares NhH  9.28e-03
     25     21    1.7    384 83  HS1152545     zs38g05.r1 Soares NhH  9.28e-03
 c   26     21    1.7    397 15  AA192801      zp93g06.s1 Stratagene  9.28e-03
     27     21    1.7    422  7  AA126160      zm79b03.r1 Stratagene  9.28e-03
     28     21    1.7    426 65  AA233020      zr46a08.r1 Soares NhH  9.28e-03
     29     21    1.7    426 83  HS1150940     zr46a08.r1 Soares NhH  9.28e-03
     30     21    1.7    429 43  G24207        human STS WI-13634.    9.28e-03
 c   31     21    1.7    445 65  AA226701      nc27g02.r1 NCI CGAP P  9.28e-03
 c   32     21    1.7    445 81  HS1146141     nc27g02.r1 NCI CGAP P  9.28e-03
 c   33     21    1.7    495 55  AA[illegible]       [illegible].r1 Stratagene  9.28e-03
 c   34     21    1.7    538 108  MM1139113     mv68c05.r1 Soares mou  9.28e-03
 c   35     21    1.7    5[illegible] 4[illegible]  AA[illegible]       [illegible].r1 Soares m[illegible]  9.28e-03
 c   36     21    1.7    546 22  AA211579      zn56b11.s1 Stratagene  9.28e-03
 c   37     21    1.7    565 60  AA208970      mw72g12.r1 Soares mou  9.28e-03
 c   38     21    1.7    565 103  MMAA81239     mw72g12.r1 Soares mou  9.28e-03
 c   39     21    1.7    565 20  AA208970      mw72g12.r1 Soares mou  9.28e-03
 c   40     21    1.7    578 13  AA205532      zq67a04.s1 Stratagene  9.28e-03
     41     21    1.7    586  7  AA123429      mp91a08.r1 Soares 2Nb  9.28e-03
     42     21    1.7    586 98  MMAA23429     mp91a08.r1 Soares 2Nb  9.28e-03
 c   43     21    1.7    617 35  AA[illegible]       mx[illegible].r1 Barstead m  9.28e-03
 c   44     21    1.7    667 11  AA129507      mt54g08.r1 Stratagene  9.28e-03
 c   45     21    1.7    667 51  AA129507      mt54g08.r1 Stratagene  9.28e-03
```

```
                              ALIGNMENTS

RESULT     1
LOCUS       G28564        2202 bp    DNA             STS       01-JUL-1996
DEFINITION  human STS SHGC-35295.
ACCESSION   G28564
NID         g1408379
KEYWORDS    STS sequence; primer; sequence tagged site.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2202)
  AUTHORS   Myers,R.M.
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Richard M. Myers
            Stanford Human Genome Center (SHGC)
            Stanford University School of Medicine
            Department of Genetics, M-344, Stanford, CA 94305, USA
            Tel: 4157259687
            Fax: 4157259689
            Email: myers@shgc.stanford.edu
            Primer A: AATTGAACTCACTGCTTTTGAGG
            Primer B: CAGCTTCAGGTATCATCTTCACC
            STS size: 206
            PCR Profile:
                    Initial incubation: 94 degrees C for 90 seconds

                    Denaturation:        94 degrees C for 15 seconds
                    Annealing:           62 degrees C for 23 seconds
                    Polymerization:      72 degrees C for 30 seconds
                    PCR Cycles:          30
                    Thermal Cycler:      Perkin Elmer 9600
            Protocol:
                    Template:          25 ng
                    Primer:            each 1 uM
                    dNTPs:             each 200 uM
                    Taq Polymerase:    0.05 units/ul
                    Total Vol:         10 ul

            Buffer:
                    MgCl2:             2.5 mM
                    KCl:               50  mM
                    Tris-HCl:          20  mM
                    pH:                8.3

            Prepared with primer pairs provided by Sand.., derived from M..960
            -- Washington University/Merck EST sequence
FEATURES             Location/Qualifiers
     source          1..2202
                     /organism="Homo sapiens"
                     /map="12"
     STS             1926..2131
     primer_bind     1926..1948
     primer_bind     complement(2109..2131)
BASE COUNT      720 a    442 c    439 g    601 t
ORIGIN

  Query Match            4.2%;  Score 53;  DB 44;  Length 2202;
  Best Local Similarity  61.9%;  Pred. No. 1.76e-55;
  Matches  138;  Conservative   0;  Mismatches  85;  Indels   0;  Gaps   0;

Db      368 ttttactaaataattttttgatctttacgtagtatgtagagacaaagcatcaggatgtca 427
             |||  |||| | || |   | |   | ||||   || ||   | ||   | ||||| |||   |||
Qy      261 AGGTTCAAAGAGGGATGCTCACTTCTGGGCTGTCTGCAGGCATTACGCATCGGGATATCA 320

Db      428 ttatggagcagtaacttgtgaaggctgcaaaggatttttaaaagaagcatccgaaaaaa 487
              |||||||        | |||||||| || || |  |||||||||||||||| | |   | |
Qy      321 CTATGGAGTCTGGTCGTGTGAAGGATGTAAGGCTTTTTTAAAAGAAGCATTCAAGGACA 380

Db      488 tttagtatattcatgtcgaggatcaaggattgtattattaatagcacacacgaaacag 547
             |    |   |||    |||| ||    ||||  | ||||   ||   ||||   |||   || || |
Qy      381 TAATGATTATATTTGTCCAGCTACAAATCAGTGTACAATCGATAAAAACGGGGCAAGAG 440

Db      548 ctgtcaatactgcaggttacagagatgtattgcgtttggaatg 590
             ||| ||   |||| |   | | || |||    |   | ||||||
Qy      441 CTGCCAGGCTTGCCGACTTCGGAAGTGTTACGAAGTGGGAATG 483


RESULT     2
ID   MM7156      standard; RNA; EST; 390 BP.
AC   W13715;
NI   g1297753
DT   30-APR-1996 (Rel. 47, Created)
DT   04-MAR-1997 (Rel. 51, Last updated, Version 2)
DE   ma83b04.r1 Soares mouse p3NMF19 5 Mus musculus cDNA clone 317263 5'
DE   similar to gb:L13740 EARLY RESPONSE PROTEIN NAK1 (HUMAN);
KW   EST.
OS   Mus musculus (house mouse)
OC   Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
OC   Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
OC   Mus.
RN   [1]
RP   1-390
RA   Marra M., Hillier L., Allen M., Bowles M., Dietrich N.,
```

```
RA   Dubuque T., Geisel S., Kucaba T., Lacy M., Le M., Martin J.,
RA   Morris M., Schellenberg K., Steptoe M., Tan F., Underwood K.,
RA   Moore B., Theising B., Wylie T., Lennon G., Soares B., Wilson R.,
RA   Waterston R.;
RT   "The WashU-HHMI Mouse EST Project";
RL   Unpublished.
CC   Contact: Marra M/Mouse EST Project WashU-HHMI Mouse EST Project
CC   Washington University School of MedicineP 4444 Forest Park Parkway,
CC   Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810
CC   Email: mouseest@watson.wustl.edu This clone is available
CC   royalty-free through LLNL ; contact the IMAGE Consortium
CC   (info@image.llnl.gov) for further information. Seq primer: ETPrimer
CC   High quality sequence stop: 385.
FH   Key             Location/Qualifiers
FH
FT   source          1..390
FT                   /organism="Mus musculus"
FT                   /note="Vector: pT7T3D (Pharmacia) with a modified
FT                   polylinker: Site_1: Not I: Site_2: Eco RI: 1st strand cDNA
FT                   was primed with a Not I - oligo(dT) primer [5'
FT                   TGTTACCAATCTGAAGTGGGAGCGGCCGCATTTTTTTTTTTTTTTTTTT 3'],
FT                   double-stranded cDNA was size selected, ligated to Eco RI
FT                   adapters (Pharmacia), digested with Not I and cloned into
FT                   the Not I and Eco RI sites of a modified pT7T3 vector
FT                   (Pharmacia). Library went through one round of
FT                   normalization to a Cot = 5. Library constructed by Bento
FT                   Soares and M.Fatima Bonaldo. RNA was kindly provided by Dr

FT                   Minoru Ko (Wayne State University)."
FT                   /clone="317263"
FT                   /clone_lib="Soares mouse p3NMF19.5"
FT                   /dev_stage="19.5 dpc total fetus"
FT                   /lab_host="DH10B (ampicillin resistant)"
FT   mRNA            <1..>390
SQ   Sequence 390 BP; 70 A; 106 C; 114 G; 100 T; 0 other;

  Query Match            4.1%;  Score 52;  DB 26;  Length 390;
  Best Local Similarity  62.7%;  Pred. No. 1.42e-53;
  Matches  128;  Conservative   0;  Mismatches  76;  Indels   0;  Gaps   0;

Db       80 tttaactatcccgacactgagacacttctgaaatctgcagtactgacatcgattcatcg 139
             ||| ||||| || ||    |   ||||||| ||| || ||||   ||| || |   || || ||
Qy      489 CTTCACCATTCCCACTTCGTAACACTTCCGAAGTCGGCAGCCTGGCAGCTCTTCCCCCC 430

Db      140 tctcttgtccactgggcagtttttatttgccaggcaaacatattttgcatttttctgcac 199
                 | || || | ||  || |   ||   | ||  | |||| ||| |      || |     || |
Qy      429 GTTTTTATCGATTGTACACTGATTTGTAGCTGGACAAATATAATCATTATGTCCTTGAAT 370

Db      200 cgttcttcttgaagaagcccttgcagcctcgcaggtgcagactcataggtgctgacagc 259
                 |||| || || ||| |||| || || || || |   | |||||| ||||| |   |   |
Qy      369 GCTTCTTTTAAAAAAGGCCTTACATCCTTCACACGACCAGACTCCATAGTGATATCCCGA 310

Db      260 agcgttgtcgccgcacacggcaca 283
              ||||  |||| ||| || || ||
Qy      309 TGCGTAATCGCTGCAGATACGCCA 286


RESULT     3
LOCUS       AA023625         593 bp    mRNA              EST       21-JAN-1997
DEFINITION  mh80b06.r1 Soares mouse placenta 4NbMP13.5 14.5 Mus musculus cDNA
            clone 457235 5' similar to gb:M12674 ESTROGEN RECEPTOR (HUMAN);
            gb:M38651 Mouse estrogen receptor mRNA, complete cds (MOUSE);.
ACCESSION   AA023625
NID         g1487542
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 593)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:274123
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 492.
FEATURES             Location/Qualifiers
     source          1..593
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGGAAATTTTTTTTTTTTTTTTTTTTT
                     T 3']; double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization, and was
                     constructed by Bento Soares and M.Fatima Bonaldo."
                     /clone="457235"
                     /clone_lib="Soares mouse placenta 4NbMP13.5 14.5"
                     /sex="unknown"
                     /dev_stage="adult"
                     /lab_host="DH10B"
     mRNA            <1..>593
BASE COUNT      144 a    165 c    119 g    165 t
ORIGIN

  Query Match             2.2%;  Score 40;  DB 28;  Length 593;
  Best Local Similarity  74.4%;  Pred. No. 1.35e-31;
  Matches   61;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Db    174 tctttgcccagttgatcatatgagccaggtccctatctgctaggttggtcaatgaggcca 233
          |||| |||||| ||||||| || |||| |||| | || || |  |||||||    |   ||
Cp    784 TCTTGGCCCAGCTGATCATGTGTACCAACCCCTTGTCGGCCAACTTGGTCAGGGACATCA 725

Db    234 tcattgaggcttcactgaaggg 255
          |:|| ||||| ||  ||||||:|
Cp    724 TCATGGAGGCCTCGGTGAAGGG 703


RESULT    4
LOCUS       N59515        280 bp    mRNA            EST       28-JAN-1997
DEFINITION  yv57g05.r1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone
            246872 5' similar to SW LPH1_MOUSE P45448 LIVER RECEPTOR HOMOLOG ;.
ACCESSION   N59515
NID         g1203405
KEYWORDS    EST.
SOURCE      human
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 280)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E.,
            Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Insert Length: 728   Std Error: 0.00
            Seq primer: T7
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1..280
                     /organism="Homo sapiens"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Pac I - oligo(dT) primer
                     [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
                     /clone="246872"
                     /clone_lib="Soares fetal liver spleen 1NFLS"
                     /sex="male"
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
     mRNA            <1..>280
BASE COUNT       80 a     45 c     75 g     61 t     19 others
ORIGIN

  Query Match             2.1%;  Score 26;  DB 25;  Length 280;
  Best Local Similarity  54.0%;  Pred. No. 3.28e-09;
  Matches   94;  Conservative    0;  Mismatches   80;  Indels    0;  Gaps    0;

Db     26 tcngggtancattangggctcctcanctgtgnaagctgnaagggattttttaaacgnaca 85
          || || || || || ||  ||       |||| | | || |||  |||||.|.  | |
Qy    310 TCGGGATATCACTATGGAGCTCAGTCGTGTGAAGGATGTAAGGGCTTTTTTAAAACAAGC 369

Db     86 gtccaaaataataaaaggtacacntgnatataaaaccagnactaccgaattgncaaaacn 145
           | |||    ||||    || |  ||    |    |    |  | ||   ||| |   |||
Qy    370 ATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCAGTGTACAATCGATAAAAAC 429

Db    146 caggaaagcgttgtcccttactgtcgttttcaaaaatgtctaagtgttggaatg 199
          | | | ||| | || |    ||| ||  |||  || |||       || ||||||
Qy    430 CGGCGAAGAGCTGCCAGGCCTTGCCGACTTCGGAAGTGTTACGAAGTGGGAATG 483


RESULT    5
LOCUS       DM36D7S       156 bp    DNA             STS       16-APR-1996
DEFINITION  D. melanogaster STS determined from European Mapping Project
            cosmid.
ACCESSION   Z70980
NID         g1263754
KEYWORDS    sequence tagged site.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Arthropoda;
            Tracheata; Insecta; Pterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 156)
  AUTHORS   European Drosophila Mapping,Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-APR-1996) Michael Ashburner, Department of Genetics,
            Downing St., Cambridge CB2 3EH, England
COMMENT     STS_name = Dm36D7S
            clone_name = 36D7
            STS_from_promoter = SP6
            vector_class = cosmid, Lorist 6
            origin_of_clone = Oregon-R
            in_situ_site_primary - 98C
            BLAST_program = BLASTN
            database_searched = EMBL
            database_version - 45.0 and updates till date_of_search
            date_of_search = 08-01-1996
            BLAST_program = BLASTX
            database_searched = SWISSPROT
            database_version - 32.0
            date_of_search = 15-12-1995.
FEATURES             Location/Qualifiers
     source          1..156
                     /organism="Drosophila melanogaster"
                     /strain="Oregon-R"
                     /clone="36D7"
BASE COUNT       28 a     37 c     30 g     27 t     34 others
ORIGIN

  Query Match              1.9%,  Score 24,  DB 36,  Length 156,
  Best Local Similarity  19.6%;  Pred. No. 1.71e-06;
  Matches     9;  Conservative   26;  Mismatches  11;  Indels   0;  Gaps   0;

Db      21 qsksrrqrrcqrkrrqqskssksrrqkkcrkskrqssqqaqtrr 66
           |: ::::| :||:::: |:::::::: : | :: : :: ||| ::
Qy     597 GCCCCGAGTGCGGGAGCTGCTGCTGGACGCCCTGAGCCCCGAGCAG 642


RESULT     6
ID   MM94013     standard; RNA; EST; 536 BP.
AC   W20940;
NI   g1297857
DT   08-MAY-1996 (Rel. 47, Created)
DT   05-MAR-1997 (Rel. 51, Last updated, Version 2)
DE   mb96d08.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone 337263 5'
DE   similar to gb:M64497 APOLIPOPROTEIN AI REGULATORY PROTEIN-1
DE   (HUMAN); gb:U07635 Mus musculus ovalbumin upstream promoter
DE   transcription factor (MOUSE):.
KW   EST.
OS   Mus musculus (house mouse)
OC   Eukaryotae, mitochondrial eukaryotes; Metazoa; Chordata;
OC   Vertebrata, Eutheria, Rodentia; Sciurognathi; Muridae; Murinae;
OC   Mus.
RN   [1]
RP   1-536
RA   Marra M., Hillier L., Allen M., Bowles M., Dietrich N
RA   Dubuque T., Geisel S., Kucaba T., Lacy M., Le M., Martin J.,
RA   Morris M., Schellenberg K., Steptoe M., Tan F., Underwood K.,
RA   Moore B., Theising B., Wylie T., Lennon G., Soares B., Wilson R.,
RA   Waterston R.;
RT   "The WashU-HHMI Mouse EST Project",
RL   Unpublished.
CC   Contact: Marra M/Mouse EST Project WashU HHMI Mouse EST Project
CC   Washington University School of MedicineP 4444 Forest Park Parkway,
CC   Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax. 314 286 1810
CC   Email: mouseest@watson.wustl.edu This clone is available
CC   royalty-free through LLNL ; contact the IMAGE Consortium
CC   (info@image.llnl.gov) for further information. MGI 218663 Seq
CC   primer: ETPrimer High quality sequence stop: 395
FH   Key             Location/Qualifiers
FH
FT   source          1..536
FT                   /organism="Mus musculus"
FT                   /note="Vector: pT7T3D (Pharmacia) with a modified
FT                   polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
FT                   was primed with a Not I - oligo(dT) primer [5'
FT                   TGTTACCAATCTGAAGTGGGAGCGGCCGCATTTTTTTTTTTTTTTTTTT 3'],
FT                   double-stranded cDNA was size selected, ligated to Eco RI
FT                   adapters (Pharmacia), digested with Not I and cloned into
FT                   the Not I and Eco RI sites of a modified pT7T3 vector
FT                   (Pharmacia). Library went through one round of
FT                   normalization to a Cot = 5. Library constructed by Bento
FT                   Soares and M.Fatima Bonaldo. RNA was kindly provided by Dr

FT                   Minoru Ko (Wayne State University)."
```

```
FT                   /clone="337263"
FT                   /clone_lib="Soares mouse p3NMF19.5"
FT                   /dev_stage="19.5 dpc total fetus"
FT                   /lab_host="DH10B (ampicillin resistant)"
FT   mRNA            <1..>536
SQ   Sequence 536 BP, 113 A, 175 C, 164 G, 84 T, 0 other,

  Query Match              1.9%;  Score 24;  DB 97;  Length 536;
  Best Local Similarity  66.7%;  Pred. No. 1.71e-06;
  Matches    48;  Conservative    0;  Mismatches   24;  Indels    0;  Gaps    0;

Db     205 [illegible] 264
           ||| | |||||  |  |||  |||||    ||||||  |||||| ||    |||||    |
Qy     412 TGTACAATCGATAAAAACCGGCGCAAGAGCTGGCAGGC[illegible]GACT[illegible]GGAAGTGTTAC 471

Db     265 aaagtgggcatg 276
            ||||||| |||
Qy     472 GAAGTGGGAATG 483


RESULT     7
ID   HS1164487  standard, RNA, EST, 326 BP.
AC   AA249798;
NI   g1881068
DT   13-MAR-1997 (Rel. 51, Created)
DT   13-MAR-1997 (Rel. 51, Last updated, Version 1)
DE   111371.seq.F Human fetal heart, Lambda ZAP Express Homo sapiens
DE   cDNA 5'.
KW   EST.
OS   Homo sapiens (human)
OC   Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
OC   Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   1-326
RA   Liew C.C.;
RT   "cDNAs from human fetal heart (1997)";
RL   Unpublished.
CC   Contact: Liew CC Molecular Cardiology University of Toronto Banting
CC   Institute, 100 College St., Toronto, Ontario, M5G1L5 Tel:
CC   4169788758 Fax: 4169788659 Email: liewcc@utcc.utoronto.ca PCR
CC   PRimers FORWARD: 5' GCCAAGCTCGAAATTAACCCTCACTAAAGGG 3' BACKWARD: 5'
CC   CCAGTGAATTGTAATACGACTCACTATAGGGCG 3' Seq primer: 5'
CC   GAAATTAACCCTCACTAAAGGG 3'.
FH   Key             Location/Qualifiers
FH
FT   source          1..326
FT                   /organism="Homo sapiens"
FT                   /note="Vector: Lambda ZAP Express, Site_1: EcoRI, Site_2:
FT                   XhoI; mRNA was purified from human fetal hearts (8-10
FT                   weeks). cDNA was synthesized using a XhoI-Oligo dT
FT                   adaptor primer. EcoRI adaptors were ligated, followed by
FT                   digestion with XhoI, fordirectional cloning into
FT                   predigested lambda ZAP Express."
FT                   /clone_lib="Human fetal heart, Lambda ZAP Express"
FT                   /lab_host="E. coli XL1-Blue"
FT   mRNA            <1..>326
SQ   Sequence 326 BP, 63 A, 87 C, 86 G, 90 T, 0 other;

  Query Match              1.8%;  Score 22;  DB 85;  Length 326;
  Best Local Similarity  78.9%;  Pred. No. 5.95e-04;
  Matches    30;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Db     257 cgtctgcagctcattatctgggcacaaggcggtaccca 294
           |||| |||  ||  | |||  ||||||||||||||||
Cp     550 CGTCGGCACTTCTCTGTCTCCGCACAAGGCGGTACCCA 513


RESULT     8
LOCUS       AA249798        326 bp    mRNA             EST       11-MAR-1997
DEFINITION  111371.seq.F Human fetal heart, Lambda ZAP Express Homo sapiens
            cDNA 5'.
ACCESSION   AA249798
NID         g1881068
```

```
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 326)
  AUTHORS   Liew,C.C.
  TITLE     cDNAs from human fetal heart (1997)
  JOURNAL   Unpublished (1997)
COMMENT
            Contact: Liew CC
            Molecular Cardiology
            University of Toronto
            Banting Institute, 100 College St., Toronto, Ontario, M5G1L5
            Tel: 4169788758
            Fax: 4169785650
            Email: liewcc@utcc.utoronto.ca
            PCR PRimers
            FORWARD: 5' GCCAAGCTCGAAATTAACCCTCACTAAAGGG 3'
            BACKWARD: 5' CCAGTGAATTGTAATACGACTCACTATAGGGCG 3'
            Seq primer: 5' GAAATTAACCCTCACTAAAGGG 3'.
FEATURES             Location/Qualifiers
     source          1..326
                     /organism="Homo sapiens"
                     /note="Vector: Lambda ZAP Express; Site_1: EcoRI; Site_2:
                     XhoI; mRNA was purified from human fetal hearts (8-10
                     weeks). cDNA was synthesized using a XhoI-Oligo dT
                     adaptor-primer. EcoRI adaptors were ligated, followed by
                     digestion with XhoI, for directional cloning into
                     predigested lambda ZAP Express."
                     /clone_lib="Human fetal heart, Lambda ZAP Express"
                     /lab_host="E. coli XL1-Blue"
     mRNA            <1..>326
BASE COUNT       63 a     87 c     86 g     90 t
ORIGIN

  Query Match             1.9%;  Score 22;  DB 69;  Length 326;
  Best Local Similarity   78.9%;  Pred. No. 5.95e-04;
  Matches    30;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Db    257 cgtctgcagctcattatctgggcacaaggcggtaccca 294
          |||:| |||  ||   | |||  ||||||||||||||||
Qy    550 CGTCGGCACTTCTCTGTCTCCGCACAAGGCGGTACCCA 513


RESULT    9
LOCUS       AA057786      382 bp    mRNA            EST       01-FEB-1997
DEFINITION  zf15e04.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone
            377022 3'.
ACCESSION   AA057786
NID         g1550437
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 382)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E.,
            Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1130   Std Error: 0.00
            Seq primer: -40M13 fwd. from Amersham
            High quality sequence stop: 272.
FEATURES             Location/Qualifiers
     source          1..382
                     /organism="Homo sapiens"
                     /note="Organ: heart; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCATCTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library constructed by
                     M.Fatima Bonaldo. This library was constructed from the
                     same fetus as the fetal lung library, Soares fetal lung
                     NbHL19W."
                     /clone="377022"
                     /clone_lib="Soares fetal heart NbHH19W"
                     /sex="unknown"
                     /dev_stage="19 weeks"
                     /lab_host="DH10B (ampicillin resistant)"
     mRNA            complement(<1..>382)
BASE COUNT       70 a    103 c    127 g     79 t      3 others
ORIGIN

  Query Match             1.9%;  Score 22;  DB 30;  Length 382;
  Best Local Similarity   78.9%;  Pred. No. 5.95e-04;
  Matches    30;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Db    145 gctgctggaggcacaggccccaggcctggagtgctca 182
          |||:||||| || | | |||||| || |  ||||||||
Qy    615 GCTGCTGGACGGCCTGAGCCCCGAGCAGCTAGTGCTCA 652


RESULT    10
LOCUS       AA021604      399 bp    mRNA            EST       31-JAN-1997
DEFINITION  ze67e04.s1 Soares retina N2b4HR Homo sapiens cDNA clone 364038 3'
            similar to gb:X01677 GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE,
            LIVER (HUMAN);.
ACCESSION   AA021604
NID         g1485265
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 399)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E.,
            Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1473   Std Error: 0.00
            Seq primer: -40M13 fwd. from Amersham
            High quality sequence stop: 369.
FEATURES             Location/Qualifiers
     source          1..399
                     /organism="Homo sapiens"
                     /note="Organ: eye; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
```

```
                        strand cDNA was primed with a Not I - oligo(dT) primer [5'
                        TGTTACCAATCTGAAGTGGGAGCGGCCGCATCTTTTTTTTTTTTTTTTTT 3'],
                        double-stranded cDNA was size selected, ligated to Eco RI
                        adapters (Pharmacia), digested with Not I and cloned into
                        the Not I and Eco RI sites of a modified pT7T3 vector
                        (Pharmacia). The retinas were obtained from a 55 year old
                        Caucasian  and total cellular poly(A)+ RNA was extracted 6
                        hrs after their removal. The retina RNA  was kindly
                        provided by Roderick R. McInnes M.D. Ph.D. from the
                        University  of Toronto. Library constructed by Bento
                        Soares and M.Fatima Bonaldo."
                        /clone="364039"
                        /clone_lib="Soares retina N2b4HR"
                        /sex="male"
                        /tissue_type="retina"
                        /dev_stage="55 year old"
                        /lab_host="DH10B (ampicillin resistant)"
     mRNA               complement(<1..>399)
BASE COUNT        74 a     92 c    132 g     98 t       3 others
ORIGIN

  Query Match             1.8%;  Score 22;  DB 21;  Length 399;
  Best Local Similarity  72.9%;  Pred. No. 5.95e-04;
  Matches   35;  Conservative    0;  Mismatches  13;  Indels    0;  Gaps    0;

Db      277 atgatcttaacaatgatgtcgattaaggacaatgccagccccagcatt 324
            ||||||||| |     ||||||  ||||| || |    |||||||| | ||
Cp      905 ATGAGCTTGCCGGGGTGGTCAATTGAGCGCCACATCAGCCCCATCATT 858


RESULT   11
ID   HSW0986    standard; RNA; EST; 424 BP.
AC   W96098;
NI   g1426005
DT   18-JUL-1996 (Rel. 48, Created)
DT   28-FEB-1997 (Rel. 51, Last updated, Version 2)
DE   ze08h09.r1 Soares fetal heart NbHH19W Homo sapiens cDNA clone
DE   358433 5' similar to gb:X52773 RETINOIC ACID RECEPTOR RXR ALPHA
DE   (HUMAN);.
KW   EST.
OS   Homo sapiens (human)
OC   Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
OC   Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   1-424
RA   Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,
RA   Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,
RA   Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,
RA   Trevaskis E., Waterston R., Williamson A., Wohldmann P., Wilson R.;
RT   "The WashU-Merck EST Project";
RL   Unpublished.
CC   Contact: Wilson RK WashU-Merck EST Project Washington University
CC   School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis,
CC   MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:
CC   est@watson.wustl.edu This clone is available royalty-free through
CC   LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for
CC   further information. Insert Length: 1285 Std Error: 0.00 Seq
CC   primer: mob.REGA+ET.
FH   Key             Location/Qualifiers
FH
FT   source          1..424
FT                   /organism="Homo sapiens"
FT                   /note="Organ: heart; Vector: pT7T3D (Pharmacia) with a
FT                   modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
FT                   strand cDNA was primed with a Not I - oligo(dT) primer [5'
FT                   TGTTACCAATCTGAAGTGGGAGCGGCCGCATCTTTTTTTTTTTTTTTTTT 3'],
FT                   double stranded cDNAwas size selected, ligated to Eco RI
FT                   adapters (Pharmacia), digested with Not I and cloned into
FT                   the Not I and Eco RI sites of a modified pT7T3 vector
FT                   (Pharmacia). Library went through one round of
FT                   normalization to a Cot - 5. Library constructed by M.Fatim
 a
FT                   Bonaldo. This library was constructed from the same fetus
```

```
FT                   as the fetal lung library, Soares fetal lung NbHL19W."
FT                   /clone="358433"
FT                   /clone_lib="Soares fetal heart NbHH19W"
FT                   /sex="unknown"
FT                   /dev_stage="19 weeks"
FT                   /lab_host="DH10B (ampicillin resistant)"
FT   mRNA            <1..>424
SQ   Sequence 424 BP; 128 A; 92 C; 117 G; 85 T; 2 other;

  Query Match             1.8%;  Score 22;  DB 89;  Length 424;
  Best Local Similarity  62.8%;  Pred. No. 5.95e-04;
  Matches   54;  Conservative    0;  Mismatches  32;  Indels    0;  Gaps    0;

Db       36 cacatctgtgccatctgtggagacagatcctcaggaaagcactacggggtatacagttgt 95
            ||| |||| ||  ||||   | ||       | || ||| | ||||| || || |      |||
Qy      280 CACTTCTCCCCTGTCTGCAGCCATTACCCATCCCCATATCACTATCCAGTCTCCTCCGTGT 339

Db       96 gaaggctgcaaagggttcttcaagag 121
            ||||| || || |   || || || ||
Qy      340 GAAGGATGTAAGGCCTTTTTTAAAAG 365


RESULT   12
LOCUS       W96098        424 bp    mRNA            EST       29-NOV-1996
DEFINITION  ze08h09.r1 Soares fetal heart NbHH19W Homo sapiens cDNA clone
            358433 5' similar to gb:X52773 RETINOIC ACID RECEPTOR RXR-ALPHA
            (HUMAN);.
ACCESSION   W96098
NID         g1426005
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 424)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1285   Std Error: 0.00
            Seq primer: mob.REGA+ET.
FEATURES             Location/Qualifiers
     source          1..424
                     /organism="Homo sapiens"
                     /note="Organ: heart; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCATCTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot - 5. Library constructed by
                     M.Fatima Bonaldo. This library was constructed from the
                     same fetus as the fetal lung library, Soares fetal lung
                     NbHL19W."
                     /clone="358433"
                     /clone_lib="Soares fetal heart NbHH19W"
                     /sex="unknown"
                     /dev_stage="19 weeks"
                     /lab_host "DH10B (ampicillin resistant)"
    mRNA             <1..>424
BASE COUNT      128 a     92 c    117 g     85 t      2 others
ORIGIN

  Query Match             1.9%;  Score 22;  DB 2;  Length 424;
  Best Local Similarity  62.9%;  Pred. No. 5.95e-04;
  Matches   54;  Conservative    0;  Mismatches  32;  Indels    0;  Gaps    0;

Db        36 cacatctgtgccatctgtggagacagatcctcaggaaagcactacggggtatacagttgt 95
             ||| |||| ||  ||||  | ||     | || ||| | ||||| || || |     |||
Qy       280 [illegible] 339

Db        96 gaaggctgcaaagggttcttcaagag 121
             ||||| || || |  || || || ||
Qy       340 GAAGGATGTAAGGCCTTTTTTAAAAG 365


RESULT   13
LOCUS        AA024507     452 bp    mRNA            EST       30-JAN-1997
DEFINITION   ze76d08.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone
             364911 3'.
ACCESSION    AA024507
NID          g1489467
KEYWORDS     EST.
SOURCE       human.
  ORGANISM   Homo sapiens
             Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
             Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 452)
  AUTHORS    Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
             Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
             Parsons,J., Rifkin,L., Pohlfing,T., Tan,F., Trevaskis,E.,
             Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE      WashU-Merck EST Project
  JOURNAL    Unpublished (1995)
COMMENT
             Contact: Wilson RK
             WashU-Merck EST Project
             Washington University School of Medicine
             4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
             Tel: 314 286 1800
             Fax: 314 286 1810
             Email: est@watson.wustl.edu
             This clone is available royalty-free through LLNL ; contact the
             IMAGE Consortium (info@image.llnl.gov) for further information.
             Insert Length: 783   Std Error: 0.00
             Seq primer: -40M13 fwd. from Amersham
             High quality sequence stop: 398.
FEATURES             Location/Qualifiers
     source          1..452
                     /organism="Homo sapiens"
                     /note="Organ: heart; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCATCTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adaptors (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library constructed by
                     M.Fatima Bonaldo. This library was constructed from the
                     same fetus as the fetal lung library, Soares fetal lung
                     NbHL19W."
                     /clone="364911"
                     /clone_lib="Soares fetal heart NbHH19W"
                     /sex="unknown"
                     /dev_stage="19 weeks"
                     /lab_host "DH10B (ampicillin resistant)"
    mRNA             complement(<1..>452)
BASE COUNT      112 a    105 c    132 g    101 t      2 others
ORIGIN
```

```
  Query Match             1.8%;  Score 22;  DB 28;  Length 452;
  Best Local Similarity  73.9%;  Pred. No. 5.95e-04;
  Matches   34;  Conservative    0;  Mismatches  12;  Indels    0;  Gaps    0;

Db       375 [illegible] 420
             ||||| | |||  ||| |||  | | |||  || ||||||||||||
Qy       591 [illegible] 636


RESULT   14
LOCUS        AA196298     480 bp    mRNA            EST       22-JAN-1997
DEFINITION   [illegible] s1 Stratagene muscle 937209 Homo sapiens cDNA clone 627954
             3'.
ACCESSION    AA196298
NID          g1791889
KEYWORDS     EST.
SOURCE       human.
  ORGANISM   Homo sapiens
             Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
             Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 480)
  AUTHORS    Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
             Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
             Parsons,J., Rifkin,L., Pohlfing,T., Tan,F., Trevaskis,E.,
             Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE      WashU-Merck EST Project
  JOURNAL    Unpublished (1995)
COMMENT
             Contact: Wilson RK
             WashU-Merck EST Project
             Washington University School of Medicine
             4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
             Tel: 314 286 1800
             Fax: 314 286 1810
             Email: est@watson.wustl.edu
             This clone is available royalty-free through LLNL ; contact the
             IMAGE Consortium (info@image.llnl.gov) for further information.
             Seq primer: -40M13 fwd. from Amersham
             High quality sequence stop: 204.
FEATURES             Location/Qualifiers
     source          1..480
                     /organism="Homo sapiens"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI; Cloned unidirectionally. Primer: Oligo dT. Skeletal
                     muscle from patient with malignant hyperthermia.  Average
                     insert size: 1.0 kb; Uni-ZAP XR Vector; ~5' adaptor
                     sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor sequence: 5'
                     CTCGAGTTTTTTTTTTTTTTTTTT 3'"
                     /clone="627954"
                     /clone_lib="Stratagene muscle 937209"
                     /dev_stage "adult"
                     /lab_host "SOLR (kanamycin resistant)"
    mRNA             complement(<1..>480)
BASE COUNT       84 a    151 c    133 g    107 t      5 others
ORIGIN

  Query Match             1.9%;  Score 23;  DB 16;  Length 480;
  Best Local Similarity  69.0%;  Pred. No. 2.38e-05;
  Matches   40;  Conservative    0;  Mismatches  18;  Indels    0;  Gaps    0;

Db       203 [illegible] 260
             |||||| ||| || |||  || | | | ||  ||| | ||||  ||||| || || | | |
Qy       250 [illegible] 193


RESULT   15
LOCUS        AA179181     498 bp    mRNA            EST       31-DEC-1996
DEFINITION   zp46b09.r1 Stratagene Hela cell s3 937216 Homo sapiens cDNA clone
             612473 5' similar to [illegible] ORF.F2F1050. ;.
ACCESSION    AA179181
NID          g1760550
KEYWORDS     EST.
SOURCE       human.
```

```
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 498)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Ritkin,L., Rohlfing,T., Tan,F., Trevaskis,E.,
            Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 302.
FEATURES             Location/Qualifiers
     source          1..498
                     /organism="Homo sapiens"
                     /note="Vector: pBluescript SK-, Site_1: EcoRI, Site_2:
                     XhoI; Cloned unidirectionally.  Primer: Oligo dT. HeLa S3
                     epithelioid carcinoma cells grown to semi-confluency
                     without induction. Average insert size: 1.5 kb; Uni-ZAP XR
                     Vector. -5' adaptor sequence: 5' GAATTCGGCACGAG 3' -3'
                     adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTTTT 3'"
                     /clone="612473"
                     /clone_lib="Stratagene HeLa cell s3 937216"
                     /sex="female"
                     /dev_stage="HeLa S3 cell line"
                     /lab_host="SOLR (kanamycin resistant)"
     mRNA            <1..>498
BASE COUNT      113 a    164 c    129 g     86 t      6 others
ORIGIN

  Query Match              1.8%;  Score 22;  DB 11;  Length 498;
  Best Local Similarity  80.6%;  Pred. No. 5.95e-04;
  Matches   29;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Db         33 atccctggccagaacaagccaaggagccaagacgag 68
              ||||||| ||||||||||    ||||| |||| |||
Cp        936 ATCCCTGTCCAGAACAAGATCTGGAGCAAAGATGAG 901


Search completed: Thu Dec 18 20:07:15 1997
Job time : 365 secs.
```

```
*************************************************************************
                         MPSRCH  (TM)
*************************************************************************

       Release 2.1D John F. Collins, Biocomputing Research Unit.
       Copyright (c) 1993, 1994, 1995  University of Edinburgh, U.K.
                 Distribution rights by IntelliGenetics, Inc.

MPsrch_nn   n a  -  n a  database search  using Smith-Waterman algorithm

Run on:          Thu Dec 18 20:51 20 1997,  MasPar time 1017.64 Seconds
                                            966 897 Million cell updates/sec
Tabular output not generated.

Title:           >US-08-826-361A-24
Description:     (1-1898) from US08826361A seq
Perfect Score:   1898
N.A. Sequence.        1 CACGAATCTTTGAGAACATT           TGAATCCAGCCTTTCCACAC 1999
    Comp                GTGCTTAGAAACTCTTGTAA         . ACTTAGGTCGGAAAGGTGTG

Scoring table:   TABLE default
                 Gap 6

Nmatch   STD :   Dbase 0;  Query 0

Searched:        707517 seqs, 256659390 bases x 2

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database:        EST-STS
                   1:EST1 2:EST2 3:EST3 4:EST4 5:EST5 6:EST6 7:EST7 8:EST8
                   9:EST9 10:EST10 11:EST11 12:EST12 13:EST13 14:EST14
                   15:EST15 16:EST16 17:EST17 18:EST18 19:EST19 20:EST20
                   21:EST21 22:EST22 23:EST23 24:EST24 25:EST25 26:EST26
                   27:EST27 28:EST28 29:EST29 30:EST30 31:EST31 32:EST32
                   33:EST33 34:EST34 35:EST35 36:EST36 37:EST37 38:EST38
                   39:EST39 40:EST40 41:EST41 42:EST42 43:EST43 44:EST44
                   45:EST45 46:EST46 47:EST47 48:EST48 49:EST49 50:EST50
                   51:EST51 52:EST52 53:EST53 54:EST54 55:EST55 56:EST56
                   57:EST57 58:EST58 59:EST59 60:EST60 61:EST61 62:EST62
                   63:EST63 64:EST64 65:EST65 66:EST66 67:EST67 68:EST68
                   69:EST69 70:EST70 71:EST71 72:EST72 73:EST73 74:EST74
                   75:EST75 76:EST76 77:EST77 78:EST78 79:EST79 80:EST80
                   81:EST81 82:EST82 83:EST83 84:EST84 85:EST85 86:EST86
                   87:EST87 88:EST88 89:EST89 90:EST90 91:EST91 92:EST92
                   93:EST93 94:EST94 95:EST95 96:EST96 97:EST97 98:EST98
                   99:EST99
Database:        EST-STS-TWO
                   100:EST100 101:EST101 102:EST102 103:EST103 104:EST104
                   105:EST105 106:EST106 107:EST107 108:EST108 109:EST109
                   110:EST110 111:EST111 112:EST112 113:EST113 114:EST114
                   115:EST115 116:EST116 117:EST117 118:EST118 119:EST119
                   120:EST120 121:EST121 122:EST122 123:EST123 124:EST124
                   125:EST125 126:EST126 127:EST127 128:EST128 129:EST129
                   130:EST130 131:EST131 132:EST132 133:EST133 134:EST134
                   135:EST135 136:EST136 137:EST137 138:EST138 139:EST139
                   140:EST140 141:EST141 142:EST142 143:EST143 144:EST144
                   145:EST145 146:EST146 147:EST147 148:EST148 149:EST149
                   150:EST150 151:EST151 152:EST152 153:EST153 154:EST154
                   155:EST155 156:EST156 157:EST157 158:EST158 159:EST159
                   160:EST160 161:EST161 162:EST162 163:EST163 164:EST164
                   165:EST165 166:EST166 167:EST167 168:EST168 169:EST169
                   170:EST170 171:EST171 172:EST172 173:EST173 174:EST174
                   175:EST175 176:EST176 177:EST177 178:EST178 179:EST179
                   180:EST180 181:EST181 182:EST182 183:EST183 184:EST184
                   185:EST185 186:EST186 187:EST187 188:EST188 189:EST189
                   190:EST190 191:EST191 192:EST192 193:EST193 194:EST194
                   195:EST195 196:EST196 197:EST197 198:EST198

Statistics:      Mean 11.616;  Variance 1.945;  scale 5.972

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                SUMMARIES
                   %
Result           Query
   No.  Score    Match Length DB  ID            Description            Pred. No.
 ------------------------------------------------------------------------------
c    1     53     2.8    454  71  H82542        ys69h11.r1 Homo sapie  4.76e-58
c    2     52     2.7    390 149  W13715        mu93b04.r1 Soares mou  5.10e-56
c    3     52     2.7    403  74  H51897        ys81d07.r1 Homo sapie  5.10e-56
     4     51     2.7    412 191  AA157597      zo44a02.r1 Stratagene  5.33e-54
     5     48     2.5    508 120  W53545        md56a06.r1 Soares mou  5.19e-48
     6     48     2.5    536 147  AA003987      mq80b08.r1 Soares mou  5.19e-48
     7     41     2.2    552 120  W53477        md45d07.r1 Soares mou  1.63e-34
     8     38     2.0    365 181  AA092184      114609.seq.F Fetal he  5.63e-29
     9     38     2.0    [illegible]  F12304    y379c04.r1 Homo sapie  5.63e-29
    10     38     2.0    604 195  AA049907      mj13e03.r1 Soares mou  5.63e-29
    11     36     1.9    566 120  W53511        md45f07.r1 Soares mou  2.20e-25
    12     32     1.7    378 141  C08482        C.elegans cDNA clone   1.76e-18
    13     29     1.5    163   8  T83788        yd65a11.r1 Homo sapie  1.34e-13
    14     29     1.5    360  89  CELK105GXF    C.elegans cDNA clone   4.88e-12
    15     29     1.5    431 137  AA044343      zk54a08.r1 Soares pre  4.88e-12
    16     29     1.5    578 125  W74889        me32d01.r1 Soares mou  4.88e-12
    17     27     1.4    368 198  N44459        SW3ICA1324SK Brugia m  1.63e-10
    18     27     1.4    483   6  T78477        yd21e03.r1 Homo sapie  1.63e-10
    19     25     1.3    120 141  C09766        C elegans cDNA clone   1.38e-07
    20     24     1.3    141  65  H28316        yl60e10.r1 Homo sapie  3.44e-06
c   21     24     1.3    202 143  N98093        K1423F Fetal heart, L  3.44e-06
    22     24     1.3    336 154  AA011708      zi03d03.r1 Soares fet  3.44e-06
c   23     24     1.3    340 154  AA011660      zi03d03.s1 Soares fet  3.44e-06
    24     25     1.3    360  89  CELK082E1F    C.elegans cDNA clone   1.38e-07
    25     25     1.3    382 118  W39474        zc76d10.r1 Pancreatic  1.38e-07
    26     25     1.3    393 146  T26966        NIB1114L12R Infant br  1.38e-07
    27     24     1.3    416  49  M78027        EST01614 Homo sapiens  3.44e-06
    28     24     1.3    466 154  AA019673      zi13f11.r1 Soares fet  3.44e-06
    29     24     1.3    487  53  H13041        ym51d05.r1 Homo sapie  3.44e-06
    30     24     1.3    530 151  W20840        mb96d08.r1 Soares mou  3.44e-06
    31     23     1.2    [illegible]  H55496    CHR220422 Homo sapien  7.64e-05
c   32     23     1.2    264  80  T30395        EST15843 Homo sapiens  7.64e-05
    33     23     1.2    300 107  HSC2GB031     H. sapiens partial cD  7.64e-05
    34     23     1.2    342  68  H74144        ys16b06.r1 Homo sapie  7.64e-05
c   35     23     1.2    344 156  AA034114      zi06d11.s1 Soares fet  7.64e-05
    36     23     1.2    374  11  [illegible]   [illegible]            7.64e-05
    37     23     1.2    425   9  T86331        yd84c11.r1 Homo sapie  7.64e-05
c   38     23     1.2    431  15  T11624        A1324F Homo sapiens c  7.64e-05
    39     23     1.2    441  18  T51819        yb54a10.s1 Homo sapie  7.64e-05
    40     23     1.2    464 165  C17508        Human placenta cDNA 5  7.64e-05
c   41     23     1.2    468 165  W49765        zc41f06.r1 Soares sen  7.64e-05
    42     23     1.2    485 173  W74443        zd75e11.s1 Soares fet  7.64e-05
c   43     23     1.2    566  76  R09448        yf19g10.s1 Homo sapie  7.64e-05
c   44     23     1.2    573 116  W20710        13e1 Human retina cDN  7.64e-05
    45     23     1.2    594 154  AA008625      mg85b11.r1 Soares mou  7.64e-05

                                ALIGNMENTS

RESULT      1
LOCUS       H82542        454 bp    mRNA               EST        09-NOV-1995
DEFINITION  ys69h11.r1 Homo sapiens cDNA clone 221121 5' similar to
            gb X51417_cds1 STEROID HORMONE RECEPTOR ERR2 (HUMAN);.
ACCESSION   H82542
NID         g1060631
KEYWORDS    EST.
```

```
SOURCE      human clone=220101 primer=M13RP1 library=Soares retina N2b4HR
            vector=pT7T3D (Pharmacia) with a modified polylinker host=DH10B
            (ampicillin resistant) Rsite1=Not I Rsite2=Eco RI 1st strand cDNA
            was primed with a Not I   oligo(dT) primer
            [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCTTTTTTTTTTTTTTTTTTT-3'],
            double-stranded cDNA was size selected, ligated to Eco RI adapters
            (Pharmacia), digested with Not I and cloned into the Not I and Eco
            RI sites of a modified pT7T3 vector (Pharmacia). The retinas were
            obtained from a 55 year old Caucasian male  and total cellular
            poly(A)+ RNA was extracted 6 hrs after their removal  The retina
            RNA was kindly provided by Roderick R. McInnes M.D  Ph.D  from the
            University of Toronto. Library constructed by Bento Soares and
            M.Fatima Bonaldo.
  ORGANISM  Homo sapiens
            Eukaryotae; Metazoa; Eumetazoa, Bilateria, Coelomata,
            Deuterostomia; Chordata; Vertebrata, Gnathostomata, Osteichthyes,
            Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
            Eutheria; Archonta; Primates; Catarrhini; Hominidae, Homo.
REFERENCE   1  (bases 1 to 454)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E , Waterston,R , Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St  Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops: 345
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL , contact the
            IMAGE Consortium (info@image llnl gov) for further information
FEATURES             Location/Qualifiers
     source          1..454
                     /organism="Homo sapiens"
                     /clone="220101"
     mRNA            <1..>454
BASE COUNT       82 a    118 c    136 g    114 t       4 others
ORIGIN

  Query Match               2.8%;  Score 53;  DB 71;  Length 454;
  Best Local Similarity  65.7%;  Pred. No. 4.76e-58;
  Matches   119;  Conservative    0;  Mismatches  61;  Indels   1;  Gaps   1;

Db      137 cttcagcatccccactttgaaggcatttcatgaaggggcaggcctggcaggacttggcgcct 196
            ||||| ||| ||||||| |    || |||     |   |||||||||||||||   ||||||||
Cp      724 CTTCACCATTCCCACTTCGTAACACTTCCGAAGTCGGCAGGCCTCCCACCTCTTGCCCCG 665

Db      197 ccgtttagtgatctcgcactcgttggtggccgggcaggtgtactcaatgttccctttggat 256
               |||   |||     ||||  || || || || ||   | || ||| | |   ||||| ||
Cp      664 GTTTTTATCGATTGTACACTGATTTGTAGCTGGACAAATATAATCATTATGTCCTTGAAT 605

Db      257 agtcctcttgaagaaggccttgcaagcctcgcaggaggccaacgcctaagtggtagccag 316
              | || || || ||||||| ||   | || || ||   |   || ||   |||| || || |
Cp      604 GCTTCTTTTAAAAAAGGCCTTACATCCTTCACACGAC-CAGACTCCATAGTGATATCCCG 546

Db      317 a 317
            |
Cp      545 A 545


RESULT      2
LOCUS       W13715          390 bp    mRNA             EST       10-SEP-1996
DEFINITION  ma83b04.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone 317263 5'
            similar to gb:I13740 EARLY RESPONSE PROTEIN NAK1 (HUMAN);
ACCESSION   W13715
NID         g1287753
```

```
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 390)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty free through LLNL , contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: ETPrimer
            High quality sequence stop: 385.
FEATURES             Location/Qualifiers
     source          1..390
                     /organism="Mus musculus"
                     /note="Vector. pT7T3D (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCATTTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library constructed by Bento
                     Soares and M Fatima Bonaldo. RNA was kindly provided by
                     Dr. Minoru Ko (Wayne State University)."
                     /clone="317263"
                     /clone_lib "Soares mouse p3NMF19.5"
                     /dev_stage="19.5 dpc total fetus"
                     /lab_host="DH10B (ampicillin resistant)"
     mRNA            <1..>390
BASE COUNT       70 a    106 c    114 g    100 t
ORIGIN

  Query Match               2.7%;  Score 52;  DB 149;  Length 390;
  Best Local Similarity  62.7%;  Pred  No  5.10e-56;
  Matches   128;  Conservative    0;  Mismatches  76;  Indels   0;  Gaps   0;

Db       80 cttaccatcccgacactgagacacttctgaaatctgcagtactgacatcagtttcggcg 139
            ||| ||||| || ||    |  ||||||| ||| || ||||   ||| || |   || || ||
Cp      724 CTTCACCATTCCCACTTCGTAACACTTCCGAAGTCGGCAGGCCTCCCACCTCTTGCCCCG 665

Db      140 tctcttgtccactgggcagttttatttgccaggcaaacatatttgcatttttctgcac 199
              | || || | ||  || |  ||   | ||  | |||| ||| |     || |    || |
Cp      664 GTTTTTATCGATTGTACACTGATTTGTAGCTGGACAAATATAATCATTATGTCCTTGAAT 605

Db      200 cgttctcttgaagaagcccttgcagccctcgcaggtgcggactccgtagtgctggcaagc 259
              |||| || || ||| |||| || || || || |   | |||||| ||||| |   |   |
Cp      604 GCTTCTTTTAAAAAAGGCCTTACATCCTTCACACGACCAGACTCCATAGTGATATCCCGA 545

Db      260 agcgttgtcgccgcacacggcaca 283
              ||||  |||| ||| || || ||
Cp      544 TGCGTAATCGCTGCAGACAGCGCA 521


RESULT      3
LOCUS       H91890          403 bp    mRNA             EST       29-NOV-1995
DEFINITION  ys81d07.r1 Homo sapiens cDNA clone 221197 5' similar to
            gb:X51417_cds1 STEROID HORMONE RECEPTOR ERR2 (HUMAN);.
```

```
ACCESSION   H91890
NID         g1087468
KEYWORDS    EST.
SOURCE      human clone=221197 primer=M13RP1 library=Soares retina N2b4HR
            vector=pT7T3D (Pharmacia) with a modified polylinker host=DH10B
            (ampicillin resistant) Rsite1=Not I Rsite2=Eco RI 1st strand cDNA
            was primed with a Not I - oligo(dT) primer
            [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCTTTTTTTTTTTTTTTTTT-3'],
            double stranded cDNA was size selected, ligated to Eco RI adapters
            (Pharmacia), digested with Not I and cloned into the Not I and Eco
            RI sites of a modified pT7T3 vector (Pharmacia). The retinas were
            obtained from a 55 year old Caucasian male  and total cellular
            poly(A)+ RNA was extracted 6 hrs after their removal. The retina
            RNA was kindly provided by Roderick R. McInnes M.D. Ph.D. from the
            University of Toronto. Library constructed by Bento Soares and
            M.Fatima Bonaldo.
  ORGANISM  Homo sapiens
            Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
            Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
            Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
            Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 403)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops: 330
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
FEATURES             Location/Qualifiers
     source          1..403
                     /organism="Homo sapiens"
                     /clone="221197"
     mRNA            <1..>403
BASE COUNT       73 a    113 c    118 g     97 t      2 others
ORIGIN

  Query Match             2.7%;  Score 52;  DB 74;  Length 403;
  Best Local Similarity  66.7%;  Pred. No. 5.10e-56;
  Matches  104;  Conservative    0;  Mismatches   52;  Indels    0;  Gaps    0;

Db     112 gttcaggatccccactttgaggcatttcatgaagcggcaggcctggcaggatttgcgcct 171
            ||||| ||| |||||||| |    || |||    |   ||||||||||||||  ||||||||
Cp     724 CTTCACCATTCCCACTTCGTAACACTTCCGAACTCCCCAGCCCTGCCAGCCTTTTCCCCG 665

Db     172 ccgtttggtgatctcgcactcgttggtggccgggcagctgtactcaatgttcccttggat 231
                |||    |||     ||||  || || || || ||   | || ||| | |  ||||| ||
Cp     664 GTTTTTATCGATTGTACACTGATTTGTAGCTGGACAAATATAATCATTATGTCCTTGAAT 605

Db     232 agtcctcttgaagaaggccttgcaagcctcgcagga 267
               | || || || |||||||| ||   | || || ||
Cp     604 GCTTCTTTTAAAAAAGGCCTTACATCCTTCACACGA 569


RESULT     4
LOCUS       AA157597      412 bp    mRNA             EST       11-DEC-1996
DEFINITION  zo44a02.r1 Stratagene endothelial cell 937223 Homo sapiens cDNA
            clone 589706 5' similar to gb:X07282 RETINOIC ACID RECEPTOR BETA-2
            (HUMAN);.
ACCESSION   AA157597
NID         g1729205
```

```
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 412)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E.,
            Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 351.
FEATURES             Location/Qualifiers
     source          1..412
                     /organism="Homo sapiens"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI. Cloned unidirectionally. Primer: Oligo dT.
                     Umbilical vein endothelial cells, passaged once. Average
                     insert size: 1.0 kb, Uni-ZAP XR Vector; ~5' adaptor
                     sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor sequence: 5'
                     CTCGAGTTTTTTTTTTTTTTTTTT 3'"
                     /clone="589706"
                     /clone_lib="Stratagene endothelial cell 937223"
                     /dev_stage="umbilical vein, 1 passage"
                     /lab_host="SOLR (kanamycin resistant)"
     mRNA            <1..>412
BASE COUNT      129 a     89 c    101 g     89 t      4 others
ORIGIN

  Query Match             2.7%;  Score 51;  DB 191;  Length 412;
  Best Local Similarity  64.9%;  Pred. No. 5.33e-54;
  Matches  111;  Conservative    0;  Mismatches   60;  Indels    0;  Gaps    0;

Db      29 gggtaccactatggggtcagcgcctgtgagggatgtaagggctttttccgcagaagtatt 88
            || || |||||||| ||| |   | ||||| |||||||||| ||||||      ||||| |||
Qy     548 GGATATCACTATGGAGTCTGGTCGTGTGAAGGATGTAAGGCCTTTTTTAAAAGAAGCATT 607

Db      89 caggaagaatatgatttacacttgtcaccgagataagaactgtgttattaataaatcacc 148
            ||      |||     ||| | |||||           ||  | |||    ||   | | |    |
Qy     608 TAAGGATATAATGATTATATTTGTCCAGCTACAAATCAGTGTACAATCGATAAAAACCCG 667

Db     149 aggaatcgatgccaatactgtcgactccagaagtgctttgaagtgggaatc 199
              | ||  | |||||    ||| ||||| | |||||| |   ||||||||||
Qy     668 CCTAAGAGCTCCCAGGCCCTCCCCATTCGGAAGTGTTACCAAGTGGCAATC 718


RESULT     5
LOCUS       W53545        508 bp    mRNA             EST       07-JUN-1996
DEFINITION  md56a06.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
            clone 372370 5' similar to gb:X57528 M.musculus mRNA for retinoic
            acid receptor-alpha (MOUSE);.
ACCESSION   W53545
NID         g1357435
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
            Murinae; Mus.
REFERENCE   1  (bases 1 to 508)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:233802
            Seq primer: mob.REGA+ET
            High quality sequence stop: 352.
FEATURES             Location/Qualifiers
     source          1..508
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGGAAATTTTTTTTTTTTTTTTTTTTTT
                     T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
                     14.5dpc embryos [total RNA provided by  Minoru Ko, Wayne
                     State Univ., from 2 ]; double-stranded cDNA was ligated to
                     Eco RI adaptors (Pharmacia), digested with Not I and
                     cloned into the Not I and Eco RI sites of the modified
                     pT7T3 vector. Library went through one round of
                     normalization, and was constructed by Bento Soares and
                     M.Fatima Bonaldo."
                     /clone="372370"
                     /clone_lib="Soares mouse embryo NbME13.5 14.5"
                     /sex="unknown"
                     /tissue_type="embryo"
                     /dev_stage="13.5-14.5dpc total fetus"
                     /lab_host="DH10B"
     mRNA            <1..>508
BASE COUNT      123 a    168 c    127 g     90 t
ORIGIN

  Query Match              2.5%;  Score 48;  DB 120;  Length 508;
  Best Local Similarity  63.6%;  Pred. No. 5.19e-48;
  Matches  112;  Conservative    0;  Mismatches  64;  Indels    0;  Gaps    0;

Db     233 catccggctaccactatgggatcagcgcctgtgagggctgtaaggccttcttccgacgaa 292
           |||| || || |||||||| ||| |  | ||||| || ||||||| ||| ||    | |||
Qy     543 CATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATGTAAGGCCTTTTTTAAAAGAA 602

Db     293 gcatccagaagaacatggtgtatacgtgtcaccgggacaagaactgcatcatcaacaagg 352
           |||| ||      | |  |  ||||  ||||         ||  | || |  ||| | ||
Qy     603 GCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCAGTGTACAATCGATAAAA 662

Db     353 tgacccggaaccgctgccagtactgccggctgcagaaatgtttcgacgtgggcatg 408
                || ||  ||||||||  |||||| || | ||| |||| ||| ||||| |||
Qy     663 ACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTACGAAGTGGGAATG 718


RESULT      6
LOCUS       AA003987        536 bp    mRNA            EST       23-JUL-1996
DEFINITION  mq80b08.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
            clone 439287 5' similar to gb:X57528 M.musculus mRNA for retinoic
            acid receptor-alpha (MOUSE);.
ACCESSION   AA003987
NID         g1447462
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
```

```
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
            Murinae; Mus.
REFERENCE   1  (bases 1 to 536)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubaque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:264623
            Seq primer: mob.REGA+ET
            High quality sequence stop: 346.
FEATURES             Location/Qualifiers
     source          1..536
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGGAAATTTTTTTTTTTTTTTTTTTTTT
                     T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
                     14.5dpc embryos [total RNA provided by  Minoru Ko, Wayne
                     State Univ., from 2 ]; double-stranded cDNA was ligated to
                     Eco RI adaptors (Pharmacia), digested with Not I and
                     cloned into the Not I and Eco RI sites of the modified
                     pT7T3 vector. Library went through one round of
                     normalization, and was constructed by Bento Soares and
                     M.Fatima Bonaldo."
                     /clone="439287"
                     /clone_lib="Soares mouse embryo NbME13.5 14.5"
                     /sex="unknown"
                     /tissue_type="embryo"
                     /dev_stage="13.5-14.5dpc total fetus"
                     /lab_host="DH10B"
     mRNA            <1..>536
BASE COUNT      129 a    180 c    131 g     96 t
ORIGIN

  Query Match              2.5%;  Score 48;  DB 147;  Length 536;
  Best Local Similarity  63.6%;  Pred. No. 5.19e-48;
  Matches  112;  Conservative    0;  Mismatches  64;  Indels    0;  Gaps    0;

Db     286 catccggctaccactatgggatcagcgcctgtgagggctgtaaggccttcttccgacgaa 345
           |||| || || |||||||| ||| |  | ||||| || ||||||| ||| ||    | |||
Qy     543 CATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATGTAAGGCCTTTTTTAAAAGAA 602

Db     346 gcatccagaagaacatggtgtatacgtgtcaccgggacaagaactgcatcatcaacaagg 405
           |||| ||      | |  |  ||||  ||||         ||  | || |  ||| | ||
Qy     603 GCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCAGTGTACAATCGATAAAA 662

Db     406 tgacccggaaccgctgccagtactgccggctgcagaaatgtttcgacgtgggcatg 461
                || ||  ||||||||  |||||| || | ||| |||| ||| ||||| |||
Qy     663 ACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTACGAACTGGGAATG 718


RESULT      7
LOCUS       W53477          552 bp    mRNA            EST       03-JUN-1996
DEFINITION  md45d07.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
            clone 371341 5' similar to gb:X07282 RETINOIC ACID RECEPTOR BETA-2
            (HUMAN);, gb:X57528 M.musculus mRNA for retinoic acid receptor-alpha
            (MOUSE);.
ACCESSION   W53477
```

```
NID         g1357321
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
            Murinae; Mus.
REFERENCE   1  (bases 1 to 552)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:232773
            Seq primer: mob REGA+ET
            High quality sequence stop: 319.
FEATURES             Location/Qualifiers
     source          1..552
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGGAAATTTTTTTTTTTTTTTTTTTTTT
                     T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
                     14.5dpc embryos [total RNA provided by  Minoru Ko, Wayne
                     State Univ., from 2 ]; double-stranded cDNA was ligated to
                     Eco RI adaptors (Pharmacia), digested with Not I and
                     cloned into the Not I and Eco RI sites of the modified
                     pT7T3 vector. Library went through one round of
                     normalization, and was constructed by Bento Soares and
                     M.Fatima Bonaldo."
                     /clone="371341"
                     /clone_lib="Soares mouse embryo NbME13.5 14.5"
                     /sex="unknown"
                     /tissue_type="embryo"
                     /dev_stage="13.5-14.5dpc total fetus"
                     /lab_host="DH10B"
     mRNA            <1..>552
BASE COUNT      140 a    151 c    169 g     92 t
ORIGIN

  Query Match             2.2%;  Score 41;  DB 120;  Length 552;
  Best Local Similarity  62.2%;  Pred. No. 1.63e-34;
  Matches  130;  Conservative    0;  Mismatches  77;  Indels    2;  Gaps    2;

Db     230 ctgcttcgtttgccaggacaagtcatcaggctaccactatgcggtcagtgcctgcgaggg 289
            ||||    || |||    ||   |   ||||||| || |||||||| ||| |   | || || ||
Qy     520 CTGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGG 579

Db     290 gtgcaaggccttttttcgcagaagtattca-gaagaacatgatctacacttgccatcgag 348
             || |||| |||||       ||||| ||||| | | |   |||||       |  ||| | |
Qy     580 ATGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTAC 639

Db     349 ataagaactgcgtcattaacaaggtcactaggaaccgatgccagtactgccgcctgcaga 408
           | ||  | ||    |   | ||   |    | ||  | |||||| |||||| || | ||
Qy     640 A AATCAGTGTACAATCGATAAAAACCGGCGCAAGAGTTGCCAGGCCTGCCGACTTTCGA 698

Db     409 agtgctttgaagtgggcatgtcaaagagt 437
           |||| |  |||||||| |||   ||| ||
Qy     699 AGTGTTACGAAGTGGGAATGGTGAAGTGT 727
```

```
RESULT     8
LOCUS       AA092184       365 bp    mRNA            EST       24-OCT-1996
DEFINITION  1146[illegible] seq F Fetal heart, Lambda ZAP Express Homo sapiens cDNA 5'.
ACCESSION   AA092184
NID         g1637173
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 365)
  AUTHORS   Liew,C.C.
  TITLE     cDNAs from fetal heart (1996)
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Liew CC
            Molecular Cardiology
            University of Toronto
            Banting Institute, 100 College St., Toronto, Ontario, M5G1L5
            Tel: 4169788758
            Fax: 4169785650
            Email: liewcc@utcc.utoronto.ca
            PCR PRimers
            FORWARD: 5' GCCAAGCTCGAAATTAACCCTCACTAAAGGG 3'
            BACKWARD: 5' CCAGTGAATTGTAATACGACTCACTATAGGGCG 3'
            Seq primer: 5' GAAATTAACCCTCACTAAAGGG 3'.
FEATURES             Location/Qualifiers
     source          1..365
                     /organism="Homo sapiens"
                     /note="Vector: lambda ZAP Express; Site_1: EcoRI; Site_2:
                     XhoI; mRNA was purified from human fetal hearts (8-10
                     weeks). cDNA was synthesized using a XhoI-Oligo dT
                     adaptor-primer. EcoRI adaptors were ligated, followed by
                     digestion with XhoI, for directional cloning into
                     predigested lambda ZAP Express."
                     /clone_lib="Fetal heart, Lambda ZAP Express"
                     /lab_host="E. coli XL1-Blue"
     mRNA            <1..>365
BASE COUNT       88 a     85 c    103 g     89 t
ORIGIN

  Query Match             2.0%;  Score 38;  DB 181;  Length 365;
  Best Local Similarity  77.1%;  Pred. No. 5.63e-29;
  Matches   54;  Conservative    0;  Mismatches  16;  Indels    0;  Gaps    0;

Db     164 cccttcagtgaagcttcgatgatgggcttactgaccaacctggcagacagggagctaatt 223
           |||||||  || || || ||||||    |   ||||||||  |||| |||| |||| ||||
Qy     938 CCCTTCACCGAGGCCTCCATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTA 997

Db     224 cacatgatca 233
           ||||||||||
Qy     998 CACATGATCA 1007
```

```
RESULT     9
LOCUS       R52304         393 bp    mRNA            EST       18-MAY-1995
DEFINITION  yg79c14.r1 Homo sapiens cDNA clone 39602 5' similar to gb:M[illegible]960
            STEROID RECEPTOR TR2 (HUMAN);.
ACCESSION   R52304
NID         g814206
KEYWORDS    EST.
SOURCE      human clone=39602 library=Soares infant brain 1NIB vector=Lafmid BA
            host=DH10B (ampicillin resistant) primer=M13RP1 Rsite1=Not I
            Rsite2=Hind III Whole brain from a 73 days post natal female. 1st
            strand cDNA was primed with a Not I - oligo(dT) primer [5'
            AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTTTT 3']; double-stranded
            cDNA was ligated to Hind III adaptors (Pharmacia), digested with
            Not I and directionally cloned into the Not I and Hind III sites of
            the Lafmid BA vector. Library went through one round of
            normalization. Library constructed by Bento Soares and M.Fatima
            Bonaldo.
```

```
  ORGANISM  Homo sapiens
            Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
            Deuterostomia; Chordata, Vertebrata; Gnathostomata, Osteichthyes;
            Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
            Eutheria. Archonta; Primates; Catarrhini; Hominidae; Homo
REFERENCE   1  (bases 1 to 393)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            GDB: G00-412-143
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops: 303
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information
FEATURES             Location/Qualifiers
     source          1..393
                     /organism="Homo sapiens"
                     /clone="39682"
BASE COUNT      122 a     88 c     84 g     97 t      2 others
ORIGIN

  Query Match             2.0%;  Score 38;  DB 26;  Length 393;
  Best Local Similarity  66.7%;  Pred. No. 5.63e-29;
  Matches   74;  Conservative    0;  Mismatches  37;  Indels    0;  Gaps    0;

Db     271 aggaccaaataaggtttttgatctttgcgtagtatgtggagacaaagcatcaggacgtca 330
            |||  |||| | || |   | |  | ||||   || ||   | ||   | ||||| |||   |||
Qy     496 AGGTTCAAAGAGGGATGCTCACTTCTGCGCTGTCTGCAGCGATTACGCATCGGGATATCA 555

Db     331 ttatggagcagtaacttgtgaaggctgcaaaggtttttttaaaagnagcat 381
             |||||||        | |||||||| || || |  |||||||||| |||||
Qy     556 CTATGGAGTCTGGTCGTGTGAAGGATGTAAGGCCTTTTTTAAAAGAAGCAT 606


RESULT     10
LOCUS       AA049907        604 bp    mRNA             EST        30-DEC-1996
DEFINITION  mj13e03.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
            clone 475996 5' similar to gb:M73779 RETINOIC ACID RECEPTOR ALPHA-1
            (HUMAN); gb:X57528 M.musculus mRNA for retinoic acid receptor-alpha
            (MOUSE);.
ACCESSION   AA049907
NID         g1529596
KEYWORDS    EST
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 604)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, Mo 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:286740
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 359.
FEATURES             Location/Qualifiers
     source          1..604
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGAAATTTTTTTTTTTTTTTTTTTT
                     T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
                     14.5dpc embryos [total RNA provided by  Minoru Ko, Wayne
                     State Univ., from 2 ], double-stranded cDNA was ligated to
                     Eco RI adaptors (Pharmacia), digested with Not I and
                     cloned into the Not I and Eco RI sites of the modified
                     pT7T3 vector. Library went through one round of
                     normalization, and was constructed by Bento Soares and
                     M.Fatima Bonaldo."
                     /clone="475996"
                     /clone_lib="Soares mouse embryo NbME13.5 14.5"
                     /sex="unknown"
                     /tissue_type="embryo"
                     /dev_stage="13.5-14.5dpc total fetus"
                     /lab_host="DH10B"
     mRNA            <1..>604
BASE COUNT      142 a    194 c    153 g    115 t
ORIGIN

  Query Match             2.0%;  Score 38;  DB 196;  Length 604;
  Best Local Similarity  60.8%;  Pred. No. 5.63e-29;
  Matches  107;  Conservative    0;  Mismatches  69;  Indels    0;  Gaps    0;

Db     287 catccggctaccactatggggtcagcgcctgtgagggctgtaagggcttcttccgacgaa 346
           |||| || || |||||||| ||| |   | ||||| || ||||||| ||| ||    | |||
Qy     543 CATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATGTAAGGCCTTTTTTAAAAGAA 602

Db     347 gcatccagaagaacatggtgtatacgtgtcaccgggacaagaactgcatcatcagcaagg 406
           |||| ||      | | |  |||| ||||          ||  | || |  ||| | ||
Qy     603 GCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCAGTGTACAATCGATAAAA 662

Db     407 tgacccggaaccggctgcaatactgccggctgcaggaatgtttcgacgttgggatg 462
                 || ||  |     |||  |||||| || | ||| |||| ||| || || |||
Qy     663 ACGGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTACGAAGTGGGAATG 718


RESULT     11
LOCUS       W53511          566 bp    mRNA             EST        03-JUN-1996
DEFINITION  md45f07.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
            clone 371365 5' similar to gb:X07282 RETINOIC ACID RECEPTOR BETA-2
            (HUMAN); gb:X57528 M.musculus mRNA for retinoic acid receptor alpha
            (MOUSE);.
ACCESSION   W53511
NID         g1357336
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
            Murinae; Mus.
REFERENCE   1  (bases 1 to 566)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
```

```
COMMENT
              Contact: Marra M/Mouse EST Project
              WashU-HHMI Mouse EST Project
              Washington University School of MedicineP
              4444 Forest Park Parkway  Box 8501  St  Louis, MO 63108
              Tel: 314 286 1800
              Fax: 314 286 1810
              Email: mouseest@watson wustl edu
              This clone is available royalty-free through LLNL ; contact the
              IMAGE Consortium (info@image.llnl.gov) for further information.
              MGI:232797
              Seq primer: mob REGA+ET
              High quality sequence stop: 360.
FEATURES             Location/Qualifiers
     source          1..566
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /note="Vector  pT7T3D-Pac (Pharmacia) with a modified
                     polylinker: Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGGAAATTTTTTTTTTTTTTTTTTTTTT
                     T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
                     14.5dpc embryos [total RNA provided by  Minoru Ko, Wayne
                     State Univ., from 2 ]; double-stranded cDNA was ligated to
                     Eco RI adaptors (Pharmacia), digested with Not I and
                     cloned into the Not I and Eco RI sites of the modified
                     pT7T3 vector  Library went through one round of
                     normalization, and was constructed by Bento Soares and
                     M.Fatima Bonaldo."
                     /clone="371365"
                     /clone_lib="Soares mouse embryo NbME13.5 14.5"
                     /sex="unknown"
                     /tissue_type="embryo"
                     /dev_stage="13 5-14 5dpc total fetus"
                     /lab_host="DH10B"
     mRNA            <1..>566
BASE COUNT      142 a    158 c    170 g     96 t
ORIGIN

  Query Match             1.9%,  Score 36,  DB 120,  Length 566;
  Best Local Similarity  62.7%;  Pred. No. 2.20e-25;
  Matches   126;  Conservative    0;  Mismatches  72;  Indels   3,  Gaps   3,

Db      242 ctgcttcgtttgccaggacaagtcatcgggctaccactatgccgtcagtgcctgcgaggg 301
            ||||    || |||    ||   |   ||||||| || |||||||| |||  |    | || || ||
Qy      520 CTGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGG 579

Db      302 gtgcaagggctttttcagcagaagtattca-gaagaacatgatctacacttgccatcgag 360
             || |||| ||||||     ||||| ||||| | | |  |||||       |  ||| | |
Qy      580 ATGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATCATTATATTGTCCAGCTAC 639

Db      361 atnagaactgcgtcattaacaaggtcactaggaacctgatgccagtactgccgcttgcag 420
             | ||  | ||     ||  | ||    |     | ||  | ||||||  |||||| ||| | |
Qy      640 A-AATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCCACTT-CCC 697

Db      421 aagtgctttgaagtgggcatg 441
            ||||| |   |||||||| |||
Qy      698 AAGTGTTACGAAGTGGGAATG 718


RESULT   12
LOCUS       C08482          378 bp    mRNA              EST       09-SEP-1996
DEFINITION  C elegans cDNA clone yk77h11 : 5' end, single read.
ACCESSION   C08482
NID         g1533553
KEYWORDS    EST(expressed sequence tag).
SOURCE      Caenorhabditis elegans (strain:N2) embryo hermaphrodite cDNA to
            mRNA  clone_lib:Yuji Kohara unpublished cDNA clone:yk77h11.
  ORGANISM  Caenorhabditis elegans
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Nematoda;
            Secernentea; Rhabditia; Rhabditida; Rhabditina; Rhabditoidea;
            Rhabditidae; Caenorhabditis.
REFERENCE   1  (bases 1 to 378)
  AUTHORS   Kohara,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-APR-1996) to the DDBJ/EMBL/GenBank databases. Yuji
            Kohara, National Institute of Genetics, Gene Library Lab; Yata
            1111, Mishima, Shizuoka 411  Japan (E-mail:ykohara@lab nig ac.jp,
            Tel.0559-81-6854, Fax:0559-81-6855)
REFERENCE   2  (sites)
  AUTHORS   Kohara,Y., Motohashi,T., Tabara,H., Watanabe,H., Sugimoto,A.,
            Sano,M., Miyata,A. and Nishigaki,A.
  TITLE     Expression map of the C.elegans genome
  JOURNAL   Unpublished (1996)
FEATURES             Location/Qualifiers
     source          1..378
                     /organism="Caenorhabditis elegans"
                     /strain="N2"
                     /clone="yk77h11"
                     /clone_lib="Yuji Kohara unpublished cDNA"
                     /dev_stage="embryo"
                     /sex="hermaphrodite"
BASE COUNT      126 a     75 c     89 g     82 t      6 others
ORIGIN

  Query Match             1.7%;  Score 32;  DB 141,  Length 378;
  Best Local Similarity  65.2%;  Pred. No. 1.76e-18;
  Matches    58;  Conservative    0;  Mismatches  31,  Indels    0;  Gaps    0;

Db      122 tttgnccagtttgnggggatacggcactgggaaagcactacggtgtaacggcttgtnat 181
            || || || || ||  | |||   |||  |||| | |||||| || ||     | | ||| |
Qy      518 TTCTGCGCTGTCTGCAGTGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAA 577

Db      182 ggatgcaagggatttnncagaagaagtat 210
            ||||| |||||  |||    | |||||| ||
Qy      578 GGATGTAAGGCCTTTTTTAAAAGAAGCAT 606


RESULT   13
LOCUS       T83788          163 bp    mRNA              EST       16-MAR-1995
DEFINITION  yd65a11.r1 Homo sapiens cDNA clone 113084 5' similar to gb.M73779
            RETINOIC ACID RECEPTOR ALPHA-1 (HUMAN);.
ACCESSION   T83788
NID         g712076
KEYWORDS    EST.
SOURCE      human clone=113084 library=Soares fetal liver spleen 1NFLS
            vector=pT7T3D (Pharmacia) with a modified polylinker host=DH10B
            (ampicillin resistant) primer=M13RP1 Rsite1=Pac I Rsite2=Eco RI
            Liver and spleen from a 20 week-post conception male fetus. 1st
            strand cDNA was primed with a Pac I - oligo(dT) primer [5'
            AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTT 3'], double-stranded
            cDNA was ligated to Eco RI adaptors (Pharmacia)  digested with Pac
            I and cloned into the Pac I and Eco RI sites of the modified pT7T3
            vector.  Library went through one round of normalization. Library
            constructed by Bento Soares and M.Fatima Bonaldo.
  ORGANISM  Homo sapiens
            Eucaryotae, Metazoa; Chordata, Vertebrata, Gnathostomata; Mammalia;
            Eutheria; Primates, Catarrhini, Hominidae, Homo.
REFERENCE   1  (bases 1 to 163)
  AUTHORS   Hillier,L., Clark,N  Dubuque,T  Elliston K  Hawkins,M ,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J , Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E , Waterston,R , Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax. 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence starts: 1
            High quality sequence stops: 1
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
FEATURES             Location/Qualifiers
     source          1..163
                     /organism="Homo sapiens"
                     /clone="113084"
BASE COUNT       50 a     41 c     49 g     23 t
ORIGIN

  Query Match              1.5%;  Score 29;  DB 8;  Length 163;
  Best Local Similarity  78.4%;  Pred. No. 1.34e-13;
  Matches    40;  Conservative    0;  Mismatches  11;  Indels   0;  Gaps   0;

Db      35 cggaaccgctgccagtactgccgactgcagaagtgctttgaagtgggcatg 85
            || ||  |||||||  |||||||||| | |||||| |  |||||||| |||
Qy     668 CGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTACGAAGTGGGAATG 718


RESULT   14
LOCUS       CELK105GXF     360 bp    mRNA             EST       13-DEC-1995
DEFINITION  C.elegans cDNA clone yk105g10 : 5' end, single read.
ACCESSION   D75525
NID         g1121318
KEYWORDS    EST(expressed sequence tag).
SOURCE      Caenorhabditis elegans (strain N2, ) (library: Yuji kohara
            unpublished cDNA) Hermaphrodite embryo embryo cDNA to mRNA
  ORGANISM  Caenorhabditis elegans
            Eukaryotae; mitochondrial eukaryotes, Metazoa; Nematoda;
            Secernentea; Rhabditia; Rhabditida; Rhabditina; Rhabditoidea;
            Rhabditidae; Caenorhabditis
REFERENCE   1  (bases 1 to 360)
  AUTHORS   Kohara,Y., Motohashi,T., Tabara,H., Sugimoto,A., Watanabe,H. and
            Nishigaki,A.
  TITLE     Toward an expression map of the C elegans genome
  JOURNAL   Unpublished (1995)
COMMENT     Submitted (23 Aug 1995) to DDBJ by:
            Yuji Kohara
            Gene Library Lab.
            National Institute of Genetics
            Yata 1111, Mishima Shizuoka
            411 Japan
            Phone: 0559-81-6854
            Fax  : 0559-81-6855
            Email:ykohara@ddbj.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1..360
                     /organism="Caenorhabditis elegans"
                     /strain="N2"
                     /dev_stage="embryo"
                     /sex="Hermaphrodite"
                     /tissue_type="embryo"
                     /clone_lib="Yuji kohara unpublished cDNA"
BASE COUNT       95 a     86 c    104 g     65 t     10 others
ORIGIN

  Query Match              1.5%;  Score 28;  DB 89;  Length 360;
  Best Local Similarity  65.5%;  Pred. No. 4.88e-12;
  Matches    57;  Conservative    0;  Mismatches  30;  Indels   0;  Gaps   0;

Db      98 tgcgctgtntgcggggacgtggctctcggaaaacactacggagtgaacgcttgcaacggg 157
           |||||||| ||| | ||     ||     ||| | ||||| |||||      | ||  | ||
Qy     521 TGCGCTGTCTGCAGCGATTACGGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGA 580

Db     158 tgcaaaggattttttaggagaagcatt 184
           || || |  |||||||  |||||||||
Qy     581 TGTAAGGCCTTTTTTAAAAGAAGCATT 607


RESULT   15
LOCUS       AA044343       431 bp    mRNA             EST       04-SEP-1996
DEFINITION  zk54a08.r1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone
```

```
            486614 5'.
ACCESSION   AA044343
NID         g1522199
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 431)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer:  28M13 rev2 from Amersham
            High quality sequence stop: 358.
FEATURES             Location/Qualifiers
     source          1..431
                     /organism="Homo sapiens"
                     /note="Organ: uterus; Vector: pT7T3-Pac; Site_1: Not I;
                     Site_2: Eco RI; 1st strand cDNA was primed with a Not I -
                     oligo(dT) primer [5'
                     AACTGGAAGAATTCGCGGCCGCTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization.  Library
                     constructed by M. Fatima Bonaldo."
                     /clone="486614"
                     /clone_lib="Soares pregnant uterus NbHPU"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="DH10B"
     mRNA            <1..>431
BASE COUNT       7[illegible] a    125 c    1[illegible] g     71 t      6 others
ORIGIN

  Query Match              1.5%;  Score 28;  DB 137;  Length 431;
  Best Local Similarity  75.9%;  Pred. No. 4.88e-12;
  Matches    41;  Conservative    0;  Mismatches  13;  Indels   0;  Gaps   0;

Db     192 ggaggccctgagccccgaggtgctagagctctcgtttcgagggactgagcccc 245
           ||| ||||||||||||||||  ||||| |||| |   | |   | | |||||||| ||
Qy     856 GGACGCCCTGAGCCCCGAGCAGCTAGTGCTCACCCTCCTGGAGGCTGAGCCGCC 909


Search completed. Thu Dec 18 21.08.23 1997
Job time : 1023 secs.
```

```
 *****************************************************************
   __  __   ____    ____   ____    ____   _   _
  |  \/  | |  _ \  / ___| |  _ \  / ___| | | | |
  | |\/| | | |_) | \___ \ | |_) || |     | |_| |
  | |  | | |  __/   ___) ||  _ < | |___  |  _  |
  |_|  |_| |_|     |____/ |_| \_\ \____| |_| |_|  (TM)

 *****************************************************************

          Release 2.1D John F. Collins, Biocomputing Research Unit
          Copyright (c) 1993, 1994, 1995  University of Edinburgh, U.K.
                   Distribution rights by IntelliGenetics, Inc.

MPsrch_nn   n.a. - n.a. database search, using Smith-Waterman algorithm

Run on:           Thu Dec 18 21:17:56 1997; MasPar time 32.29 Seconds
                                             879.865 Million cell updates/sec
Tabular output not generated

Title:            >US-08-826-361A-24
Description:      (1-1898) from US08826361A seq
Perfect Score:    1898
N.A. Sequence:       1 CACGAATCTTTGAGAACATT..........TGAATCGAGCGTTTCCACAC 1898
     Comp              GTGCTTAGAAACTCTTGTAA          ACTTAGCTCGCAAAGGTGTG

Scoring table:    TABLE default
                  Gap 6

Nmatch  STD :     Dbase 0; Query 0

Searched:         81563 seqs, 21620618 bases x 2

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database:         n-issued
                    1:backl 2:51 3:52 4:53 5:54 6:55 7:56 8:PCT90 9:PCT91
                    10:PCT92 11:PCT93 12:PCT94 13:PCT95 14:PCT96

Statistics:       Mean 9.166; Variance 4.522; scale 2.027

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                   SUMMARIES
                 %
Result          Query
   No.  Score   Match Length DB  ID              Description               Pred. No.
---------------------------------------------------------------------------------
     1    197    10.4   4963  5  US-08-076-      Sequence 16, Applicati   1.78e-138
     2    197    10.4   4963  7  US-08-260-      Sequence 9, Applicati    1.78e-138
c    3     71     3.7   3469  6  US-08-312-      Sequence 1, Applicatio   2.67e-35
     4     63     3.3   6244  5  US-08-076-      Sequence 15, Applicati   3.66e-29
     5     63     3.3   6244  7  US-08-260-      Sequence 8, Applicatio   3.66e-29
c    6     62     3.3   7218  7  US-08-232       Sequence 14, Applicati   2.11e-28
     7     61     3.2   7218  7  US-08-232-      Sequence 14, Applicati   1.20e-27
     8     55     2.9   2989  1  5223606-1       Patent No. 5223606       3.80e-23
     9     53     2.8   3014  7  US-08-629-      Sequence 1, Applicatio   1.15e-21
    10     51     2.7   1576  1  5260432-1       Patent No. 5260432       3.40e-20
    11     47     2.5    387  6  US-08-561-      Sequence 1, Applicatio   2.73e-17
    12     47     2.5    387  5  US-08-087-      Sequence 1, Applicatio   2.73e-17
    13     42     2.2   2329 11  PCT-US93-0      Sequence 2, Applicati    9.81e-14
    14     42     2.2   2940  1  5171671-1       Patent No. 5171671.      9.81e-14
    15     42     2.2   3036 10  PCT-US92-0      Sequence 1, Applicatio   9.81e-14
    16     38     2.0   1662  9  PCT-US91-0      Sequence 5, Applicatio   5.76e-11
    17     35     1.8    215  6  US-08-238-      Sequence 5, Applicatio   6.09e-09
c   18     33     1.7    215  6  US-08-238-      Sequence 5, Applicatio   1.27e-07
    19     33     1.7   1952  6  US-08-333-      Sequence 1, Applicatio   1.27e-07
    20     30     1.6   2095  9  PCT-US91-0      Sequence 3, Applicatio   1.08e-05
    21     30     1.6   2130  5  US-07-952-      Sequence 1, Applicatio   1.08e-05
    22     29     1.5   1866  9  PCT-US91-0      Sequence 1, Applicatio   4.57e-05
    23     28     1.5   2204  5  US-07-952-      Sequence 2, Applicatio   1.90e-04
    24     28     1.5   1219  9  PCT-US91-0      Sequence 7, Applicati    1.90e-04
    25     28     1.5   2468  6  US-08-333-      Sequence 11, Applicati   1.90e-04
    26     26     1.4    242  7  US-08-273-      Sequence 1, Applicatio   3.07e-03
    27     26     1.4   1649 12  PCT-US94-0      Sequence 1, Applicatio   3.07e-03
    28     26     1.4   1659  6  US-08-333-      Sequence 7, Applicatio   3.07e-03
    29     26     1.4   2204  7  US-08-464-      Sequence 1, Applicati    3.07e-03
c   30     24     1.3     69  7  US-08-471-      Sequence 142, Applicat   4.51e-02
    31     24     1.3     74 13  PCT-US95-1      Sequence 100, Applicat   4.51e-02
    32     24     1.3     74 13  PCT-US95-1      Sequence 94, Applicati   4.51e-02
c   33     25     1.3     74 13  PCT-US95-1      Sequence 94, Applicati   1.19e-02
c   34     24     1.3     74 13  PCT-US95-1      Sequence 100, Applicat   4.51e-02
    35     25     1.3     75 13  PCT-US95-1      Sequence 99, Applicati   1.19e-02
c   36     24     1.3     75 13  PCT-US95-1      Sequence 99, Applicati   4.51e-02
    37     25     1.3     81 13  PCT-US95-1      Sequence 92, Applicati   1.19e-02
    38     24     1.3     81 13  PCT-US95-1      Sequence 98, Applicati   4.51e-02
c   39     25     1.3     81 13  PCT-US95-1      Sequence 92, Applicati   1.19e-02
c   40     24     1.3     81 13  PCT-US95-1      Sequence 98, Applicati   4.51e-02
    41     25     1.3     82 13  PCT-US95-1      Sequence 97, Applicati   1.19e-02
c   42     25     1.3     82 13  PCT-US95-1      Sequence 97, Applicati   1.19e-02
    43     24     1.3    105  5  US-07-865-      Sequence 13, Applicati   4.51e-02
    44     24     1.3   2009  6  US-08-273-      Sequence 4, Applicatio   4.51e-02
    45     24     1.3   3284  1  5258289-3       Patent No. 5258289.      4.51e-02

                                   ALIGNMENTS

RESULT     1
ID   US-08-076-726-16 STANDARD; DNA; UNC; 4963 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 16, Application US/08076726.
CC   Sequence 16, Application US/08076726
CC   Patent No. 5464758
CC    GENERAL INFORMATION:
CC      APPLICANT:  Gossen, Manfred
CC      APPLICANT:  Bujard, Hermann
CC      TITLE OF INVENTION:  Tight Control of Gene Expression in
CC      TITLE OF INVENTION:  Eucaryotic Cells by Tetracycline-responsive Promot
  ers
CC      NUMBER OF SEQUENCES:  16
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  Sterne, Kessler, Goldstein & Fox
CC        STREET:  1100 New York Avenue, N.W.
CC        CITY:  Washington
CC        STATE:  District of Columbia
CC        COUNTRY:  United States of America
CC        ZIP:  20005-3934
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  US/08/076,726
CC        FILING DATE:  14-JUN-1993
CC        CLASSIFICATION:  435
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  Esmond, Robert W.
CC        REGISTRATION NUMBER:  32,893
CC        REFERENCE/DOCKET NUMBER:  1442.24[illegible]01
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  (202)371-2600
CC        TELEFAX:  (202)371-2540
CC        TELEX:  248636 SSK
CC     INFORMATION FOR SEQ ID NO:  16:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  4963 base pairs
CC        TYPE:  nucleic acid
CC        STRANDEDNESS:  both
CC         TOPOLOGY: both
SQ   Sequence 4963 BP; 1259 A; 1300 C; 1278 G; 1126 T; 0 other;

  Query Match             10.4%;  Score 197;  DB 5;  Length 4963;
  Best Local Similarity   66.7%;  Pred. No. 1.78e-138;
  Matches  424;  Conservative    0;  Mismatches 209;  Indels    3;  Gaps    2;

Db    1471 CCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGCTGGTT 1530
           |||||||  || || || ||||||   |  ||||||||  |||| |||| |||| |||| 
Qy     938 CCCTTCACCGAGGCCTCCATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTA 997

Db    1531 CACATGATCAACTGGGCGAAGAGGGTGCCAGGCTTTGTGGATTTGACCCTCCATGATCAG 1590
           |||||||||| |||||| |||| | | || |||||||||||  | | |||    || || 
Qy     998 CACATGATCAGCTGGGCCAAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAA 1057

Db    1591 GTCCACCTTCTAGAATGTGCCTGGCTAGAGATCCTGATGATTGGTCTCGTCTGGCGCTCC 1650
           || |  ||  | ||  |    ||| | ||| |   | ||||| || ||   | ||||||||
Qy    1058 GTGCGGCTCTTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCA 1117

Db    1651 ATGGAGCACCCAGTGAAGCTACTGTTTGCTCCTAACTTGCTCTTGGACAGGAACCAGGGA 1710
           || || ||||| |  |||||  | ||||||||  |   |   |  |||||||| |   ||||
Qy    1118 ATTGACCACCCGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAGGATGAGGGG 1177

Db    1711 AAATGTGTAGAGGGCATGGTGGAGATCTTCGACATGCTGCTGGCTACATCATCTCGGTTC 1770
           ||||| ||||| || ||  |||| ||||| |||||||| ||||| ||   | ||  ||||
Qy    1178 AAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTT 1237

Db    1771 CGCATGATGAATCTGCAGGGAGAGGAGTTTGTGTGCCTCAAATCTATTATTTTGCTTAAT 1830
           ||    | | || || ||      | || | | | ||  ||||   | || ||  |||| |||
Qy    1238 CGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAAT 1297

Db    1831 TCTGGAGTGTACACATTTCTGTCCAGCACCCTGAAGTCTCTGGAAGAGAAGGACCATATC 1890
           ||  |  | |||| | |  |   |    |||| | |  ||     |  ||  |||   |  |
Qy    1298 TCCAGTATGTACCCTCTGGTCACAGCGACCCAGGATGCTGAC-AGCAGCGGAAGCTGGC 1356

Db    1891 CACCGAGTCCTGGACAAGATCACAGACACTTTGATCCACCTGATGGCCAAGGCAGGCCTG 1950
              |  |  ||| ||     | || ||  ||||| |      |||| ||||||    ||| |
Qy    1357 TCACTTG--CTGAACGCCGTGACCGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATC 1414

Db    1951 ACCCTGCAGCAGCAGCACCAGCGGCTGGCCCAGCTCCTCCTCATCCTCTCCCACATCAGG 2010
             ||    ||||||||    |   ||| |||||   | |||||   |   |||| |||||| |||||
Qy    1415 TCCTCCCAGCAGCAATCCATGCGCCTGGCTAACCTCCTGATGCTCCTGTCCCACGTCAGG 1474

Db    2011 CACATGAGTAACAAAGGCATGGAGCATCTGTACAGCATGAAGTGCAAGAACGTGGTGCCC 2070
           ||    ||||||||| |||||||| ||||||   || |||||||||||| || ||||| ||
Qy    1475 CATGCGAGTAACAAGGGCATGGAACATCTGCTCAACATGAAGTGCAAAAATGTGGTCCCA 1534

Db    2071 CTCTATGACCTGCTGCTGGAGATGCTGGACGCCCAC 2106
            | ||||||||||||||||||||||||| | ||||||
Qy    1535 GTGTATGACCTGCTGCTGGAGATGCTGAATGCCCAC 1570


RESULT      2
ID   US-08-260-452-9 STANDARD; DNA; UNC; 4963 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 9, Application US/08260452
CC   Sequence 9, Application US/08260452
CC   Patent No. 5650298
CC    GENERAL INFORMATION:
CC       APPLICANT: Gossen, Manfred
CC       APPLICANT: Bujard, Hermann
CC       APPLICANT: Salfeld, Jochen
CC       APPLICANT: Voss, Jeffrey
CC       TITLE OF INVENTION. Tight Control of Gene Expression in Eucaryotic
CC       TITLE OF INVENTION: Cells by Tetracycline-responsive Promoters
CC       NUMBER OF SEQUENCES: 10
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE: Lahive & Cockfield
CC         STREET: 60 State Street
CC         CITY: Boston
CC         STATE: Massachusetts
CC         COUNTRY: USA
CC         ZIP: 02109-1875
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE: Floppy disk
CC         COMPUTER: IBM PC compatible
CC         OPERATING SYSTEM: PC-DOS/MS-DOS
CC         SOFTWARE. ASCII text
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER: US/08/260,452
CC         FILING DATE:
CC         CLASSIFICATION: 435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER: 08/075,327
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME: Giulio A DeConti, Jr.
CC         REGISTRATION NUMBER: 31,503
CC         REFERENCE/DOCKET NUMBER: BBI-013CP
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE: (617) 227-7400
CC         TELEFAX: (617) 227-5941
CC     INFORMATION FOR SEQ ID NO: 9:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH: 4963 base pairs
CC         TYPE: nucleic acid
CC         STRANDEDNESS: double
CC         TOPOLOGY: circular
CC       MOLECULE TYPE: DNA (genomic)
CC       ORIGINAL SOURCE:
CC         ORGANISM: Human cytomegalovirus
CC       IMMEDIATE SOURCE:
CC         CLONE: pUHD BGR4
SQ   Sequence 4963 BP; 1259 A; 1300 C; 1278 G; 1126 T; 0 other;

  Query Match             10.4%;  Score 197;  DB 7;  Length 4963;
  Best Local Similarity   66.7%;  Pred. No. 1.78e-138;
  Matches  424;  Conservative    0;  Mismatches 209;  Indels    3;  Gaps    2;

Db    1471 CCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGCTGGTT 1530
           |||||||  || || || ||||||   |  ||||||||  |||| |||| |||| |||| 
Qy     938 CCCTTCACCGAGGCCTCCATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTA 997

Db    1531 CACATGATCAACTGGGCGAAGAGGGTGCCAGGCTTTGTGGATTTGACCCTCCATGATCAG 1590
           |||||||||| |||||| |||| | | || |||||||||||  | | |||    || || 
Qy     998 CACATGATCAGCTGGGCCAAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAA 1057

Db    1591 GTCCACCTTCTAGAATGTGCCTGGCTAGAGATCCTGATGATTGGTCTCGTCTGGCGCTCC 1650
           || |  ||  | ||  |    ||| | ||| |   | ||||| || ||   | ||||||||
Qy    1058 GTGCGGCTCTTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCA 1117

Db    1651 ATGGAGCACCCAGTGAAGCTACTGTTTGCTCCTAACTTGCTCTTGGACAGGAACCAGGGA 1710
           || || ||||| |  |||||  | ||||||||  |   |   |  |||||||| |   ||||
Qy    1118 ATTGACCACCCGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAGGATGAGGGG 1177

Db    1711 AAATGTGTAGAGGGCATGGTGGAGATCTTCGACATGCTGCTGGCTACATCATCTCGGTTC 1770
           ||||| ||||| || ||  |||| ||||| |||||||| ||||| ||   | ||  ||||
Qy    1178 AAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAACGTTT 1237

Db    1771 CGCATGATGAATCTGCAGGGAGAGGAGTTTGTGTGCCTCAAATCTATTATTTTGCTTAAT 1830
           ||    | | || || ||      | || | | | ||  ||||   | || ||  |||| |||
Qy    1238 CGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAAT 1297

Db    1831 TCTGGAGTGTACACATTTCTGTCCAGCACCCTGAAGTCTCTGGAAGAGAAGGACCATATC 1890
           ||  |  | |||| | |  |   |    |||| | |  ||     |  ||  |||   |  |
Qy    1298 TCCAGTATGTACCCTCTGGTCACAGCGACCCAGGATGCTGAC-AGCAGCGGAAGCTGGC 1356

Db    1891 CACCGAGTCCTGGACAAGATCACAGACACTTTGATCCACCTGATGGCCAAGGCAGGCCTG 1950
              |  |  ||| ||     | || ||  ||||| |      |||| ||||||    ||| |
Qy    1357 TCACTTG--CTGAACGCCGTGACCGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATC 1414

Db    1951 ACCCTGCAGCAGCAGCACCAGCGGCTGGCCCAGCTCCTCCTCATCCTCTCCCACATCAGG 2010
             ||    ||||||||    |   ||| |||||   | |||||   |   |||| |||||| |||||
Qy    1415 TCCTCCCAGCAGCAATCCATGCGCCTGGCTAACCTCCTGATGCTCCTGTCCCACGTCAGG 1474

Db    2011 CACATGAGTAACAAAGGCATGGAGCATCTGTACAGCATGAAGTGCAAGAACGTGGTGCCC 2070
```

```
              ||     |||||||||| |||||||| ||||||    || |||||||||||| || ||||| ||
Qy    1475 CATGGGAGTAACAAGGGCATGGAACATCTGCTCAACATGAAGTGCAAAAATGTGGTCCCA 1534

Db    2071 CTCTATGACCTGCTGCTGGAGATGCTGGACGCCCAC 2106
              | ||||||||||||||||||||||||| | ||||||
Qy    1535 GTGTATGACCTGCTGCTGGAGATGCTGAATGCCCAC 1570


RESULT    3
ID    US-08-312-312A-1 STANDARD; DNA; UNC; 3460 BP.
AC    xxxxxx
DT    01-JAN-1900
DE    Sequence 1, Application US/08312312A.
CC    Sequence 1, Application US/08312312A
CC    Patent No. 5550024
CC     GENERAL INFORMATION:
CC       APPLICANT:  Rothschild, Max F.
CC       APPLICANT:  Tuggle, Christopher K.
CC       APPLICANT.  Jacobson, Carol D.
CC       APPLICANT:  Mileham, Alan J.
CC       APPLICANT:  Plastow, Graham S.
CC       APPLICANT:  Vaske, David A.
CC       TITLE OF INVENTION:  Genetic Markers for Pig Litter Size
CC       NUMBER OF SEQUENCES:  9
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Zarley, McKee, Thomte, Voorhees & Sease
CC         STREET:  801 Grand Suite 3200
CC         CITY:  Des Moines
CC         STATE:  Iowa
CC         COUNTRY:  United States
CC         ZIP:  50309
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE.  PatentIn Release #1.0, Version #1.30
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER   US/08/312,312A
CC         FILING DATE:  26-SEP-1994
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 07/961,919
CC         FILING DATE:  19-OCT-1992
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER   US 07/687,708
CC         FILING DATE:  19-APR-1991
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Nebel, Heidi S.
CC         REGISTRATION NUMBER:  37.719
CC         REFERENCE/DOCKET NUMBER:  BRDC-4
CC       TELECOMMUNICATION INFORMATION
CC         TELEPHONE:  515-288-3667
CC         TELEFAX:  515-288-1338
CC     INFORMATION FOR SEQ ID NO:  1:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  3460 base pairs
CC         TYPE:  nucleic acid
CC         STRANDEDNESS:  double
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  DNA (genomic)
CC       HYPOTHETICAL:  NO
CC       ANTI-SENSE:  NO
CC       FEATURE:
CC         NAME/KEY:  misc_feature
CC         LOCATION:  303..304
CC         OTHER INFORMATION:  /note= "Gap in DNA sequence of
CC         OTHER INFORMATION:  undetermined length"
CC       FEATURE:
CC         NAME/KEY:  misc_feature
CC         LOCATION:  1197..1198
CC         OTHER INFORMATION:  /note= "Gap in DNA sequence of
CC         OTHER INFORMATION:  undetermined length"
CC       FEATURE:
CC         NAME/KEY:  exon
CC         LOCATION:  534..600
SQ    Sequence 3460 BP; 1052 A; 765 C; 634 G; 1009 T; 0 other;

  Query Match              3.7%;  Score 71;  DB 6;  Length 3460;
  Best Local Similarity  83.3%;  Pred. No. 2.67e-35;
  Matches   95;  Conservative    0;  Mismatches  18;  Indels   1;  Gaps   1;

Db     490 TTCATCATGCCCACTTCGTAGCACTTGCGTAGCCGGCAGGC-TGACAGCTCTTCCTCCTG 548
           |||| ||| |||||||||||| ||||| || || ||||||| || |||||||| | || |
Cp     723 TTCACCATTCCCACTTCGTAACACTTCCGAAGTCGGCAGGCCTGGCAGCTCTTGCGGCGG 664

Db     549 TTCTTATCAATTGTGCACTGGTTGGTAGCTGGACACATGTAGTCATTATGTCCT 602
           || ||||| ||||| ||||| || ||||||||||| || || ||||||||||||
Cp     663 TTTTTATCGATTGTACACTGATTTGTAGCTGGACAAATATAATCATTATGTCCT 610


RESULT    4
ID    US-08-076-726-15 STANDARD; DNA; UNC; 6244 BP.
AC    xxxxxx
DT    01-JAN-1900
DE    Sequence 15, Application US/08076726.
CC    Sequence 15, Application US/08076726
CC    Patent No. 5464758
CC     GENERAL INFORMATION:
CC       APPLICANT:  Gossen, Manfred
CC       APPLICANT.  Bujard, Hermann
CC       TITLE OF INVENTION.  Tight Control of Gene Expression in
CC       TITLE OF INVENTION:  Eucaryotic Cells by Tetracycline-responsive Promot
  ers
CC       NUMBER OF SEQUENCES:  16
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Sterne, Kessler, Goldstein & Fox
CC         STREET:  1100 New York Avenue, N.W.
CC         CITY:  Washington
CC         STATE:  District of Columbia
CC         COUNTRY:  United States of America
CC         ZIP.  20005-3934
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE.  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM.  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER   US/08/076,726
CC         FILING DATE:  14-JUN-1993
CC         CLASSIFICATION.  435
CC       ATTORNEY/AGENT INFORMATION.
CC         NAME:  Esmond, Robert W.
CC         REGISTRATION NUMBER:  32,893
CC         REFERENCE/DOCKET NUMBER:  0942.2490001
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (202)371-2600
CC         TELEFAX:  (202)371-2540
CC         TELEX:  248636 SSK
CC     INFORMATION FOR SEQ ID NO:  15:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH   6244 base pairs
CC         TYPE:  nucleic acid
CC         STRANDEDNESS:  both
CC         TOPOLOGY:  both
SQ    Sequence 6244 BP; 1493 A; 1749 C; 1600 G; 1402 T; 0 other;

  Query Match              3.3%;  Score 63;  DB 5;  Length 6244;
  Best Local Similarity  66.3%;  Pred. No. 3.66e-29;
  Matches  128;  Conservative    0;  Mismatches  65;  Indels   0;  Gaps   0;

Db    2365 TCTGTGGGGATGAAGCATCAGGCTGTCATTATGGTGTCCTCACCTGTGGCACCTCTAAGG 2424
           ||||   | ||| | ||||| || | ||| ||||| |||     | ||||   | ||||||
Qy     528 TCTGCAGGGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATGTAAGG 587

Db    2425 TCTTCTTTAAAAGGGCAATGGAAGGGCAGCATAACTATTTATGTGCTGGAAGAAATGACT 2484
            ||| ||||||||     ||  |||| ||   || | ||| | ||| | |   |||| | |
Qy     588 CCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCACT 647

Db    2485 GCATTGTTGATAAAATCCGGAGGAAAAACTGCCCGGCGTGTCGCCTTAGAAAGTGCTGTC 2544
           | |   | ||||||| |||  | || | ||||| ||| || || ||| | ||||| |
Qy     648 GTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTACG 707

Db    2545 AAGCTGGCATGGT 2557
           |||  || |||||
Qy     708 AAGTGGGAATGGT 720


RESULT     5
ID   US-08-260-452-8 STANDARD; DNA; UNC; 6244 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 8, Application US/08260452.
CC   Sequence 8, Application US/08260452
CC   Patent No. 5650298
CC    GENERAL INFORMATION:
CC      APPLICANT:  Gossen, Manfred
CC      APPLICANT:  Bujard, Hermann
CC      APPLICANT:  Salfeld, Jochen
CC      APPLICANT:  Voss, Jeffrey
CC      TITLE OF INVENTION:  Tight Control of Gene Expression in Eucaryotic
CC      TITLE OF INVENTION:  Cells by Tetracycline-responsive Promoters
CC      NUMBER OF SEQUENCES:  10
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  Lahive & Cockfield
CC        STREET:  60 State Street
CC        CITY:  Boston
CC        STATE:  Massachusetts
CC        COUNTRY:  USA
CC        ZIP:  02109-1875
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  ASCII text
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  US/08/260,452
CC        FILING DATE:
CC        CLASSIFICATION:  435
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  08/076,327
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  Giulio A. DeConti, Jr.
CC        REGISTRATION NUMBER:  31,503
CC        REFERENCE/DOCKET NUMBER:  BBI-013CP
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  (617) 227-7400
CC        TELEFAX:  (617) 227-5941
CC    INFORMATION FOR SEQ ID NO:  8:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  6244 base pairs
CC        TYPE:  nucleic acid
CC        STRANDEDNESS:  double
CC        TOPOLOGY:  circular
CC      MOLECULE TYPE:  DNA (genomic)
CC      ORIGINAL SOURCE:
CC        ORGANISM:  Human cytomegalovirus
CC        STRAIN:  Towne (hCMV)
CC      IMMEDIATE SOURCE:
CC        CLONE:  pUHD BGR3
SQ   Sequence 6244 BP; 1483 A; 1749 C; 1610 G; 1402 T; 0 other;

  Query Match             3.3%;  Score 63;  DB 7;  Length 6244;
  Best Local Similarity   66.3%;  Pred. No. 3.66e-29;
  Matches  128;  Conservative    0;  Mismatches   65;  Indels    0;  Gaps    0;

Db    2365 TCTCTGGGGATGAAGCATCAGGCTGTCATTATGGTGTCCTTACCTGTGGGAGTTGTAAGG 2424
           ||||  | ||| | ||||| || | ||| ||||| |||      | ||||    | |||||||
Qy     528 TCTGCAGCGATTACGCATCGGGATATCACTATGGATCTCGGTCTCTGAAGGATGTAAGC 587

Db    2425 TCTTCTTTAAAAGGGCAATGGAAGGGCAGCATAACTATTTATGTGCTGGAAGAAATGACT 2484
            ||| ||||||||    ||  |||| ||  || | ||| | ||| |    | |||| |||
Qy     588 CCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCACT 647

Db    2485 GCATTGTTGATAAAATCCGGAGGAAAAACTGCCCGGCGTGTCGCCTTAGAAAGTGCTGTC 2544
           | |   | ||||||| |||  | || | ||||| ||| || || ||| | ||||| |
Qy     648 GTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTACG 707

Db    2545 AACCTGGCATGGT 2557
           |||  || |||||
Qy     708 AAGTGGGAATGGT 720


RESULT     6
ID   US-08-232-463-14 STANDARD; DNA; UNC; 7218 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 14, Application US/08232463.
CC   Sequence 14, Application US/08232463
CC   Patent No. 5670367
CC    GENERAL INFORMATION:
CC      APPLICANT:  DORNER, F.
CC      APPLICANT:  SCHEIFLINGER, F.
CC      APPLICANT:  FALKNER, F. G.
CC      TITLE OF INVENTION:  RECOMBINANT FOWLPOX VIRUS
CC      NUMBER OF SEQUENCES:  52
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  Foley & Lardner
CC        STREET:  1800 Diagonal Road, Suite 500
CC        CITY:  Alexandria
CC        STATE:  VA
CC        COUNTRY:  USA
CC        ZIP:  22313-0299
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  US/08/232,463
CC        FILING DATE:
CC        CLASSIFICATION:  435
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US/07/935,313
CC        FILING DATE:
CC        APPLICATION NUMBER:  EP 91 114 300.6
CC        FILING DATE:  26-AUG-1991
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  BENT, Stephen A.
CC        REGISTRATION NUMBER:  29,768
CC        REFERENCE/DOCKET NUMBER:  30472/114 IMMU
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  (703)836-9300
CC        TELEFAX:  (703)683-4109
CC        TELEX:  899149
CC    INFORMATION FOR SEQ ID NO:  14:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  7218 base pairs
CC        TYPE:  nucleic acid
CC        STRANDEDNESS:  single
CC        TOPOLOGY:  linear
CC      IMMEDIATE SOURCE:
CC        CLONE:  pTZgpt-F1s
SQ   Sequence 7218 BP; 1944 A; 1491 C; 1486 G; 1929 T; 368 other;

  Query Match             3.3%;  Score 62;  DB 7;  Length 7218;
  Best Local Similarity   1.4%;  Pred. No. 2.11e-28;
  Matches    5;  Conservative  205;  Mismatches  148;  Indels    0;  Gaps    0;

Db    1083 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1142
           ::::: : :::     : : :  :: :: ::::::::  :::: ::   : :   : :  ::
Qy     729 [illegible]CACAGTGCAGCTG 788
```

```
Db   1143 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1202
            ::: ::   : :::                               :::  :   :
Qy    787 CTCGTCGGCACTTTCTCTGTCTCCGCACAAGGCGGTACCCATATCTCTCTCTCCGGAGCC 728

Db   1203 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1262
             : :::: :: ::::: :::: :   : :::::     ::   :   :::   : ::::: :
Qy    727 ACACTTTACTATTTCTACTTTCGTAACACTTTCGAATTCTCTACTCTTGGTACTCTTTCTC 668

Db   1263 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1322
           ... .    .            ... .  . . .  :  .. :: : ....:
Qy    667 CCCCTTTTTATCCATTGTACATTCATTTGTAGCTGGACAAATATAATCATTATGCTCTTTC 608

Db   1323 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1382
             : ::::::::         :::: : ..... .  .  ..   :::: :  .  . ...
Qy    607 AATCCTTCTTTTAAAAAAGGCCTTACATCCTTCACACGACCAGACTCCATAGTGATATCC 548

Db   1383 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYGTACCA 1440
           :  : : :  .. .. :   :  :  .       . .........::   :::| ||||
Qy    547 CCATCCGTAATCCTTCCAGACAGCCCACAAGTGAGCATCCCTCTTTGAATCTGGACCA 490


RESULT    7
ID   US-08-232-463 14 STANDARD, DNA, UNC, 7218 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 14, Application US/08232463.
CC   Sequence 14, Application US/08232463
CC   Patent No. 5670367
CC    GENERAL INFORMATION:
CC       APPLICANT:  DORNER, F.
CC       APPLICANT:  SCHEIFLINGER, F.
CC       APPLICANT:  FALKNER, F. G.
CC       TITLE OF INVENTION:  RECOMBINANT FOWLPOX VIRUS
CC       NUMBER OF SEQUENCES:  52
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Foley & Lardner
CC         STREET:  1800 Diagonal Road, Suite 500
CC         CITY:  Alexandria
CC         STATE:  VA
CC         COUNTRY:  USA
CC         ZIP:  22313-0299
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/232,463
CC         FILING DATE:
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/935,313
CC         FILING DATE:
CC         APPLICATION NUMBER:  EP 91 114 300.6
CC         FILING DATE:  26-AUG-1991
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  BENT, Stephen A.
CC         REGISTRATION NUMBER:  29,768
CC         REFERENCE/DOCKET NUMBER:  30472/114 IMMU
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (703)836-9300
CC         TELEFAX:  (703)683-4109
CC         TELEX:  899149
CC     INFORMATION FOR SEQ ID NO:  14:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  7218 base pairs
CC         TYPE:  nucleic acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       IMMEDIATE SOURCE:
CC         CLONE:  pTZgpt-F1s
SQ   Sequence 7218 BP; 1944 A; 1491 C; 1486 G; 1929 T; 368 other;
```

```
  Query Match             3.2%;  Score 61;  DB 7;  Length 7218;
  Best Local Similarity   0.9%;  Pred. No. 1.2e-27;
  Matches    3;  Conservative  200;  Mismatches 142;  Indels    0;  Gaps    0;

Db   1093 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1152
          ::::: : :::::::::: :      : ::::::::  :: :: ::::       : :    : :
Qy     28 GCTTTGTGCCTCTTCTTGCAACCTGTTTTCTCAGCTGTTATCTCAAGACATGGATATAAA 87

Db   1153 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1212
           ...  .         .. .:. .  .          .  :. .  ..  ... .... .:
Qy     88 AAATTCACCATCTAGCCTTAATTCTCCTTCCTCCTACAACTGCAGTCAATCCATCTTACC 147

Db   1213 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1272
          :::     : :  :::: : : : : :::::::::: : :    :   :: :: :     : :::
Qy    148 CCTGGAGCACGGCTCCATATACATACCTTCCTCCTATGTAGACAGCCACCATGAATATCC 207

Db   1273 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1332
              .. .  : ....  .   .... .. :   .    .. :   . :::::   :   : :: ::
Qy    208 ACCCATGACATTTTATAGCCCTCCTGTGATGAATTACAGCATTCCCAGCAATCTCACTAA 267

Db   1333 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1392
          :::       .   ...  ..   .    .. .    ...    : : :: :   ::   : :::
Qy    268 CTTGGAAGGTGGAGCTGGTCGCCAGACCACAAGCCCAAATGTGTTGTGGCCAACACCTGG 327

Db   1393 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYGTA 1437
            : .:..........    .    .:: :: ::   :: :: : :::|||
Qy    328 GTACTTTTCTCTTTAGTGGTCCATCGCCAGTTATCACATCTGTA 372


RESULT    8
ID   5223606-1 STANDARD; DNA, UNC, 3238 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Patent No. 5223606.
CC  Patent No. 5223606
CC      APPLICANT:  BLAUDIN DE THE, HUGHES;MARCHIO, AGNES;TIOLLAIS,
CC  PIERRE;DEJEAN, ANNE
CC      TITLE OF INVENTION: STEROID/THYROID HORMONE RECEPTOR-RELATED
CC  PROTEIN INAPPROPRIATELY EXPRESSED IN HUMAN HEPATOCELLULAR CARCINOMA
CC      NUMBER OF SEQUENCES: 11
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER  US/07/134,130
CC        FILING DATE. 17-DEC-1987
CC      PRIOR APPLICATION DATA:
CC  SEQ ID NO.1.
CC        LENGTH. 2989
SQ   Sequence 3238 BP; 900 A; 648 C; 600 G; 841 T, 249 other;

  Query Match             2.9%;  Score 55;  DB 1;  Length 2989;
  Best Local Similarity  65.0%,  Pred. No. 3.80e-23,
  Matches  130;  Conservative    0,  Mismatches  69,  Indels    1;  Gaps    1;

Db    560 CTGCTTCGTCTGCCAGGACAAATCATCAGGGTACCACTATGGGGTCAGCGCCTGTGAGGG 619
          ||||    ||||||   ||   |    |||| || || ||||||||| |||  |    |  |||| ||
Qy    520 CTGCGCTGTCTGCAGCGATTACCCATCGGGATATCACTATGGAGTCTGCTCCTGTGAAGC 579

Db    620 ATGTAAGGGCTTTTTCCGCAGAAGTATTCAGAAGAATATGATTTACACTTGTCACCGAGA 679
          |||||||| ||||||      ||||| |||||       |||      ||| | |||||
Qy    580 ATGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTAC 639

Db    680 TAAGAACTGTGTTATTAATAAAAGTCACCAGGAATCGATGCCAATACTGTCCAGTCCAGA 739
            ||  | |||   ||  ||||||  |   | | ||  | |||||   ||| ||||| | ||
Qy    640 AAATCAGTGTACAATCGATAAAAACCGGC-GCAAGAGCTGCCAGGCCTGCCGACTTCGGA 698

Db    740 AGTGCTTTGAAGTGGGAATG 759
          |||| |   ||||||||||||
Qy    699 AGTGTTACGAAGTGGGAATG 718


RESULT    9
ID   US-08-629-939 1 STANDARD; DNA; UNC, 3014 BP.
AC   xxxxxx
DT   01-JAN-1900
```

```
DE   Sequence 1, Application US/08629939
CC   Sequence 1, Application US/08629939
CC   Patent No. 5645995
CC    GENERAL INFORMATION:
CC       APPLICANT:  Kieback, Dirk G
CC       TITLE OF INVENTION:  METHODS FOR DIAGNOSING AN INCREASED
CC       TITLE OF INVENTION:  RISK OF BREAST OR OVARIAN CANCER
CC       NUMBER OF SEQUENCES:  14
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  SUGHRUE, MION, ZINN, MACPEAK & SEAS
CC         STREET:  2100 Pennsylvania Avenue, N.W., Suite 800
CC         CITY:  Washington, D.C.
CC         STATE:  D.C.
CC         COUNTRY:  U.S.A.
CC         ZIP:  20037
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/629,939
CC         FILING DATE:  12-APRIL-1996
CC         CLASSIFICATION:  435
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  KIT, Gordon
CC         REGISTRATION NUMBER:  30,764
CC         REFERENCE/DOCKET NUMBER:  A-6612
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (202) 293-7060
CC         TELEFAX:  (202) 293-7860
CC     INFORMATION FOR SEQ ID NO:  1:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  3014 base pairs
CC         TYPE:  nucleic acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  cDNA
CC       HYPOTHETICAL:  NO
SQ   Sequence 3014 BP; 623 A; 939 C; 827 G; 625 T; 0 other;

  Query Match             2.8%;  Score 53;  DB 7;  Length 3014;
  Best Local Similarity  63.7%;  Pred. No. 1.15e-21;
  Matches  123;  Conservative    0;  Mismatches  70;  Indels    0;  Gaps    0;

Db   1881 TCTGTGGGGATGAAGCATCAGGCTGTCATTATGGTGTCCTTACCTGTGGGAGCTGTAAGG 1940
          ||||  | ||| | ||||| || | ||| ||||| |||    | ||||   | |||||||
Qy    528 TCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATGTAAGG 587

Db   1941 TCTTCTTTAAGAGGGCAATGGAAGGGCAGCACAACTACTTATGTGCTGGAAGAAATGACT 2000
           ||| ||||| ||    ||  |||| ||  |  | ||  | |||  | |  | |||| | |
Qy    588 CCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCAGT 647

Db   2001 GCATCGTTGATAAAATCCGCAGAAAAAACTGCCCAGCATGTCGCCTTAGAAAGTGCTGTG 2060
          | |   | ||||||| |||  | || | |||||  || || || ||| | ||||| |  |
Qy    648 GTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTACG 707

Db   2061 AGGCTGGCATGGT 2073
          | |  || |||||
Qy    708 AAGTGGGAATGGT 720


RESULT   10
ID   5260432-1 STANDARD; DNA; UNC; 1707 BP
AC   xxxxxx
DT   01-JAN-1900
DE   Patent No. 5260432.
CC  Patent No. 5260432
CC       APPLICANT: TAKAKU, FUMIMARO; ISHIKAWA, TAKASHI; IMAWARI, MICHIO;
CC  EVANS, RONALD M.; UMESONO, KAZUHIKO
CC       TITLE OF INVENTION: HUMAN GAMMA RETINOIC ACID RECEPTOR DNA
CC       NUMBER OF SEQUENCES: 4
CC       CURRENT APPLICATION DATA:
```

```
CC          APPLICATION NUMBER:  US/07/370,407
CC          FILING DATE: 22-JUN-1989
CC   SEQ ID NO:1:
CC          LENGTH: 1576
SQ   Sequence 1707 BP; 346 A; 478 C; 467 G; 285 T; 131 other;

  Query Match             2.7%;  Score 51;  DB 1;  Length 1576;
  Best Local Similarity  65.8%;  Pred. No. 3.40e-20;
  Matches  127;  Conservative    0;  Mismatches  64;  Indels    2;  Gaps    2;

Db    473 GTGTGCAATGACAAGTCCTCTGGCTACCACTATGGGGTCAGCTCCTGTGAAGGCTGCAAG 532
          || ||||  ||  |  | || || || |||||||| ||| | || |||||||| || |||
Qy    527 GTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATGTAAG 586

Db    533 GGCTTCTTTCGCCGAAGCATCCAGAAGAACATGGTGTACACGTGTCACCGCGACAAAA-A 591
          | ||| |||    ||||||| ||         |  |  || |  |||| | || |||||  |
Qy    587 GCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTC-CAGCTACAAATCA 645

Db    592 CTGTATCATCAACAAGGTGACCAGGAATCGGTGCCAGTACTGCCGGCTACAGAAGTGCTT 651
           ||||  ||| | ||        | ||  | ||||||  |||||| | ||||||| |
Qy    646 GTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTA 705

Db    652 CGAAGTGGGCATG 664
          ||||||||| |||
Qy    706 CGAAGTGGGAATG 718


RESULT   11
ID   US-08-561-072-1 STANDARD; DNA; UNC; 387 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 1, Application US/08561072.
CC   Sequence 1, Application US/08561072
CC   Patent No. 5571791
CC     GENERAL INFORMATION:
CC       APPLICANT:  E. Brad Thompson
CC       APPLICANT:  Lynne V. Nazareth
CC       TITLE OF INVENTION:  ANTI-TUMOR ACTIVITY OF A MODIFIED FRAGMENT
CC       TITLE OF INVENTION:  OF THE GLUCOCORTICOID RECEPTOR
CC       NUMBER OF SEQUENCES:  3
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Arnold, White & Durkee
CC         STREET:  P.O. Box 4433
CC         CITY:  Houston
CC         STATE:  Texas
CC         COUNTRY:  USA
CC         ZIP:  77210
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  WordPerfect 5.1
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/561,072
CC         FILING DATE:  20-NOV-1995
CC         CLASSIFICATION:  530
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  08/087,151
CC         FILING DATE:  July 7, 1993
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Hodgins, Daniel S.
CC         REGISTRATION NUMBER:  31,026
CC         REFERENCE/DOCKET NUMBER:  UTSG:141/HOD
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  512-320-7200
CC         TELEFAX:  512-474-7577
CC         TELEX:  No. 5571791 Applicable
CC     INFORMATION FOR SEQ ID NO:  1:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  387 base pairs
CC         TYPE:  nucleic acid
CC         STRANDEDNESS:  double
CC         TOPOLOGY:  linear
CC        MOLECULE TYPE:  genomic DNA
SQ   Sequence 387 BP; 109 A; 92 C; 88 G; 98 T; 0 other;

  Query Match               2.5%;  Score 47;  DB 6;  Length 387;
  Best Local Similarity  63.0%;  Pred. No. 2.73e-17;
  Matches  126;  Conservative    0;  Mismatches  73;  Indels   1;  Gaps   1;

Db     188 TCTGCCTGGTGTGCTCTGATGAAGCTTCAGGATGTCATTATGGAGTCTTAACTTGTGGAA 247
           |||||   || |||    ||| | || || |||| ||| ||||||||||    | |||| |
Qy     519 TCTGGGTGGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAG 578

Db     248 GCTGTAAAGTTTTCTTCAAAAGAGCAGTGGAAGGACAGCACAATTACCTATGTGCTGGAA 307
           | |||||| |   || || ||||||     |   |||||||  |   ||||   | ||| | |  |
Qy     579 GATGTAAGCCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTA 638

Db     308 GGAATGATTGCATCATCGGATAAAATT-CGAAAAAACTGCCCAGCATGCCGCTATCGAA 366
              |||| | || |   |||| | |||     || || | |||||  || |||||    ||| |
Qy     639 CAAATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGA 698

Db     367 AATGTCTTCAGGCTGGAATG 386
           | |||    | |  ||||||
Qy     699 AGTGTTACGAAGTGGGAATG 718

RESULT   12
ID   US-08-087-151-1 STANDARD; DNA; UNC; 387 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 1, Application US/08087151.
CC   Sequence 1, Application US/08087151
CC   Patent No. 5468624
CC    GENERAL INFORMATION:
CC      APPLICANT:  E. Brad Thompson
CC      APPLICANT:  Lynne V. Nazareth
CC      TITLE OF INVENTION:  ANTI-TUMOR ACTIVITY OF A MODIFIED FRAGMENT
CC      TITLE OF INVENTION:  OF THE GLUCOCORTICOID RECEPTOR
CC      NUMBER OF SEQUENCES:  3
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  Arnold, White & Durkee
CC        STREET:  P.O. Box 4433
CC        CITY:  Houston
CC        STATE:  Texas
CC        COUNTRY:  USA
CC        ZIP:  77210
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  WordPerfect 5.1
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  US/08/087,151
CC        FILING DATE:  July 7, 1993
CC        CLASSIFICATION:  435
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  Hodgins, Daniel S.
CC        REGISTRATION NUMBER:  31,026
CC        REFERENCE/DOCKET NUMBER:  UTSG:141/HOD
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  512-320-7200
CC        TELEFAX:  512-474-7577
CC        TELEX:  No. 5468624 Applicable
CC     INFORMATION FOR SEQ ID NO:  1:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  387 base pairs
CC         TYPE:  nucleic acid
CC         STRANDEDNESS:  double
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  genomic DNA
SQ   Sequence 387 BP; 109 A; 92 C; 88 G; 98 T; 0 other;

  Query Match               2.5%;  Score 47;  DB 5;  Length 387;
  Best Local Similarity  63.0%;  Pred. No. 2.73e-17;
  Matches  126;  Conservative    0;  Mismatches  73;  Indels   1;  Gaps   1;

Db     188 TCTGCCTGGTGTGCTCTGATGAAGCTTCAGGATGTCATTATGGAGTCTTAACTTGTGGAA 247
           |||||   || |||    ||| | || || |||| ||| ||||||||||    | |||| |
Qy     519 TCTGGGTGGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAG 578

Db     248 GCTGTAAAGTTTTCTTCAAAAGAGCAGTGGAAGGACAGCACAATTACCTATGTGCTGGAA 307
           | |||||| |   || || ||||||     |   |||||||  |   ||||   | ||| | |  |
Qy     579 GATGTAAGCCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTA 638

Db     308 GGAATGATTGCATCATCGGATAAAATT-CGAAAAAACTGCCCAGCATGCCGCTATCGAA 366
              |||| | || |   |||| | |||     || || | |||||  || |||||    ||| |
Qy     639 CAAATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGA 698

Db     367 AATGTCTTCAGGCTGGAATG 386
           | |||    | |  ||||||
Qy     699 AGTGTTACGAAGTGGGAATG 718

RESULT   13
ID   PCT-US92-02320A-3 STANDARD; DNA; UNC; 2928 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 3, Application PC/TUS9202320A.
CC   Sequence 3, Application PC/TUS9202320A
CC    GENERAL INFORMATION:
CC      APPLICANT:  Sloan-Kettering Institute, For Cancer Research
CC      TITLE OF INVENTION:  METHODS FOR DETECTION AND TREATMENT OF CANCER
CC      NUMBER OF SEQUENCES:  6
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  COOPER & DUNHAM
CC        STREET:  30 ROCKEFELLER PLAZA
CC        CITY:  NEW YORK
CC        STATE:  NEW YORK
CC        COUNTRY:  U.S.A.
CC        ZIP:  10112
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  PatentIn Release #1.24
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  PCT/US92/02320A
CC        FILING DATE:
CC        CLASSIFICATION:
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US 675,084
CC        FILING DATE:  22-MAR-1991
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US 673,838
CC        FILING DATE:  22-MAR-1991
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  WHITE, JOHN P
CC        REGISTRATION NUMBER:  28,678
CC        REFERENCE/DOCKET NUMBER:  38094-PCT
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  (212) 977-9550
CC        TELEFAX:  (212) 664-0525
CC        TELEX:  (212) 422523 COOP UI
CC     INFORMATION FOR SEQ ID NO:  3:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  2928 base pairs
CC         TYPE:  nucleic acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  cDNA
CC       IMMEDIATE SOURCE:
CC         CLONE:  hRAR ALPHA
CC       FEATURE:
CC         NAME/KEY:  CDS
CC         LOCATION:  103..1488
CC         OTHER INFORMATION:
SQ   Sequence 2928 BP; 582 A; 1011 C; 789 G; 546 T; 0 other;
```

```
  Query Match             2.2%;  Score 42;  DB 10;  Length 2928;
  Best Local Similarity   60.6%;  Pred. No  9.81e-14;
  Matches  120;  Conservative    0;  Mismatches  78;  Indels   0;  Gaps   0;

Db     364 TGCTTTGTCTGTCAGGACAAGTCCTCAGGCTACCACTATGGGGTCAGCGCCTGTGAGGGC 423
           |||  ||||||    ||  |  | || || || |||||||| ||| |  | ||||| ||
Qy     521 TGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGA 580

Db     424 TGCAAGGGCTTCTTCCGCCGCAGCATCCAGAAGAACATGGTGTACACGTGTCACCGGGAC 483
           || |||| ||| ||     | ||||| ||     | |  |  || |  ||||
Qy     581 TGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACA 640

Db     484 AAGAACTGCATCATCAACAAGGTGACCCGGAACCGCTGCCAGTACTGCCGACTGCAGAAG 543
           ||  | || |  ||| | ||       || ||  ||||||||  ||||||||| | ||||
Qy     641 AATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAG 700

Db     544 TGCTTTGAAGTGGGCATG 561
           || |  |||||||| |||
Qy     701 TGTTACGAAGTGGGAATG 718


RESULT   14
ID   5171671-1 STANDARD; DNA; UNC; 3185 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Patent No. 5171671.
CC  Patent No. 5171671
CC       APPLICANT: EVANS, RONALD M.;ONG, ESTELITA S.;SEGUI,
CC  PRUDIMAR S.;THOMPSON, CATHERINE C.;UEMSONO, KAZUHIKO
CC  GUGUERE, VINCENT
CC       TITLE OF INVENTION: RETINOIC ACID RECEPTOR COMPOSITION
CC       NUMBER OF SEQUENCES: 2
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER: US/07/546,256
CC         FILING DATE: 06-AUG-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER: 276,536
CC         FILING DATE: 30-NOV-1988
CC         APPLICATION NUMBER: 128,331
CC         FILING DATE: 02 DEC 1987
CC  SEQ ID NO:1:
CC         LENGTH: 2940
SQ   Sequence 3185 BP; 594 A; 1010 C; 790 G; 546 T; 245 other;

  Query Match             2.2%;  Score 42;  DB 1;  Length 2940;
  Best Local Similarity   60.6%;  Pred. No  9.81e-14;
  Matches  120;  Conservative    0;  Mismatches  78;  Indels   0;  Gaps   0;

Db     364 TGCTTTGTCTGTCAGGACAAGTCCTCAGGCTACCACTATGGGGTCAGCGCCTGTGAGGGC 423
           |||  ||||||    ||  |  | || || || |||||||| ||| |  | ||||| ||
Qy     521 TGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGA 580

Db     424 TGCAAGGGCTTCTTCCGCCGCAGCATCCAGAAGAACATGGTGTACACGTGTCACCGGGAC 483
           || |||| ||| ||     | ||||| ||     | |  |  || |  ||||
Qy     581 TGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACA 640

Db     484 AAGAACTGCATCATCAACAAGGTGACCCGGAACCGCTGCCAGTACTGCCGACTGCAGAAG 543
           ||  | || |  ||| | ||       || ||  ||||||||  ||||||||| | ||||
Qy     641 AATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAG 700

Db     544 TGCTTTGAAGTGGGCATG 561
           || |  |||||||| |||
Qy     701 TGTTACGAAGTGGGAATG 718


RESULT   15
ID   PCT-US92-02320A-1 STANDARD; DNA; UNC; 3036 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 1, Application PC/TUS9202320A
CC   Sequence 1, Application PC/TUS9202320A
CC    GENERAL INFORMATION:
CC       APPLICANT:  Sloan-Kettering Institute, For Cancer Research
CC       TITLE OF INVENTION:  METHODS FOR DETECTION AND TREATMENT OF CANCER
CC       NUMBER OF SEQUENCES:  6
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  COOPER & DUNHAM
CC         STREET:  30 ROCKEFELLER PLAZA
CC         CITY:  NEW YORK
CC         STATE:  NEW YORK
CC         COUNTRY:  U.S.A.
CC         ZIP:  10112
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.24
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US92/02320A
CC         FILING DATE:
CC         CLASSIFICATION:
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 675,084
CC         FILING DATE:  22-MAR-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 673,838
CC         FILING DATE:  22-MAR-1991
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  WHITE, JOHN P
CC         REGISTRATION NUMBER:  28,678
CC         REFERENCE/DOCKET NUMBER:  38694-PCT
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 977-9550
CC         TELEFAX:  (212) 644-0525
CC         TELEX:  (212) 422523 COOP UI
CC     INFORMATION FOR SEQ ID NO:  1:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  3036 base pairs
CC         TYPE:  nucleic acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  cDNA
CC       IMMEDIATE SOURCE:
CC         CLONE:  MYL-RAR
CC       FEATURE:
CC         NAME/KEY:  CDS
CC         LOCATION:  67..2457
CC         OTHER INFORMATION:
SQ   Sequence 3036 BP; 665 A; 1018 C; 830 G; 523 T; 0 other;

  Query Match             2.2%;  Score 42;  DB 10;  Length 3036;
  Best Local Similarity   60.6%;  Pred. No  9.81e-14;
  Matches  120;  Conservative    0;  Mismatches  78;  Indels   0;  Gaps   0;

Db    1333 TGCTTTGTCTGTCAGGACAAGTCCTCAGGCTACCACTATGGGGTCAGCGCCTGTGAGGGC 1392
           |||  ||||||    ||  |  | || || || |||||||| ||| |  | ||||| ||
Qy     521 TGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGA 580

Db    1393 TGCAAGGGCTTCTTCCGCCGCAGCATCCAGAAGAACATGGTGTACACGTGTCACCGGGAC 1452
           || |||| ||| ||     | ||||| ||     | |  |  || |  ||||
Qy     581 TGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACA 640

Db    1453 AAGAACTGCATCATCAACAAGGTGACCCGGAACCGCTGCCAGTACTGCCGACTGCAGAAG 1512
           ||  | || |  ||| | ||       || ||  ||||||||  ||||||||| | ||||
Qy     641 AATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAG 700

Db    1513 TGCTTTGAAGTGGGCATG 1530
           || |  |||||||| |||
Qy     701 TGTTACGAAGTGGGAATG 718


Search completed: Thu Dec 18 21:28:55 1997
Job time : 659 secs.
```

```
*******************************************************************

                    MPSRCH                       (TM)

*******************************************************************

        Release 2.1D John F. Collins, Biocomputing Research Unit
        Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
                Distribution rights by IntelliGenetics, Inc.

MPsrch_nn   n.a. - n.a. database search, using Smith-Waterman algorithm

Run on:          Thu Dec 19 10:15:22 1997   MasPar time 63.09 Seconds
                                            857.364 Million cell updates/sec
Tabular output not generated

Title:           >US-08-826-361A-2
Description:     (1-1251) from US08826361A.seq
Perfect Score    1251
N.A. Sequence.        1 ATGAATTACAGCATTCCCAG  . . .   ACGTCAGGCATGCGAGGTGA 1251
     Comp:              TACTTAATGTCGTAAGGGTC    .     TGCAGTCCGTACGCTCCACT

Scoring table:   TABLE default
                 Gap 6

Nmatch   STD :   Dbase 0;  Query 0

Searched:        81563 seqs  21620619 bases x 2

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database:        n-issued
                   1:backl 2:51 3:52 4:53 5:54 6:55 7:56 8:pct90 9:pct91
                   10:pct92 11:pct93 12:pct94 13:pct95 14:pct96

Statistics:      Mean 8.788;  Variance 4.452;  scale 1.974

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                 SUMMARIES
                   %
Result           Query
   No.   Score   Match Length DB  ID             Description             Pred  No
 -------------------------------------------------------------------------------
      1    135    10.8   4963  5  US-08-076-   Sequence 16, Applicati  2.62e-85
      2    135    10.8   4963  7  US-08-260-   Sequence 9, Applicatio  2.62e-85
c     3     71     5.7   2460  6  US-08-312-   Sequence 1, Applicati   5.59e-35
      4     63     5.0   6244  5  US-08-076-   Sequence 15, Applicati  5.80e-29
      5     63     5.0   6244  7  US-08-260-   Sequence 8, Applicati   5.80e-29
c     6     62     5.0   7219  7  US-08-232-   Sequence 14, Applicati  3.22e-28
      7     55     4.4   2989  1  5223606-1    Patent No. 5223606      4.64e-23
      8     53     4.2   3014  7  US-08-629-   Sequence 1, Applicati   1.32e-21
      9     51     4.1   1576  1  5260432-1    Patent No. 5260432      3.68e-20
     10     47     3.8    387  5  US-08-087-   Sequence 1, Applicatio  2.64e-17
     11     47     3.8    387  6  US-08-561-   Sequence 1, Applicatio  2.64e-17
     12     42     3.4   2928 10  PCT-US92-0   Sequence 3, Applicati   8.32e-14
     13     42     3.4   2940  1  5171671-1    Patent No. 5171671      8.32e-14
     14     42     3.4   3036 10  PCT-US92-0   Sequence 1, Applicatio  8.32e-14
     15     38     3.0   1662  9  PCT-US91-0   Sequence 5, Applicatio  4.46e-11
     16     35     2.8    215  6  US-08-238-   Sequence 5, Applicatio  4.44e-09
     17     33     2.6   1952  6  US-08-333-   Sequence 1, Applicatio  8.95e-08
c    18     30     2.4    215  6  US-08-238-   Sequence 5, Applicatio  7.26e-06
     19     30     2.4   2095  9  PCT-US91-0   Sequence 3, Applicatio  7.26e-06
```

```
     20     30     2.4   2130  5  US-07-952-   Sequence 1, Applicati   7.26e-06
     21     29     2.3   1866  9  PCT-US91-0   Sequence 1, Applicati   3.04e-05
     22     28     2.2   2204  5  US-07-952-   Sequence 3, Applicatio  1.25e-04
     23     28     2.2   2219  9  PCT-US91-0   Sequence 7, Applicati   1.25e-04
     24     28     2.2   2469  6  US-08-333-   Sequence 11, Applicati  1.25e-04
     25     28     2.2   7218  7  US-08-232-   Sequence 14, Applicati  1.25e-04
     26     26     2.1   1649 12  PCT-US94-0   Sequence 1, Applicati   1.99e-03
     27     26     2.1   1659  6  US-08-333-   Sequence 7, Applicatio  1.99e-03
     28     26     2.1   2304  7  US-08-464-   Sequence 1, Applicatio  1.99e-03
c    29     25     2.0     74 13  PCT-US95-1   Sequence 94, Applicati  7.68e-03
     30     25     2.0     75 13  PCT-US95-1   Sequence 99, Applicati  7.68e-03
c    31     25     2.0     81 13  PCT-US95-1   Sequence 92, Applicati  7.68e-03
     32     25     2.0     92 13  PCT-US95-1   Sequence 97, Applicati  7.68e-03
c    33     25     2.0     81 13  PCT-US95-1   Sequence 97, Applicat   7.68e-03
c    34     24     1.9     63  7  US-08-471-   Sequence 142, Applicat  2.89e-02
     35     24     1.9     74 13  PCT-US95-1   Sequence 94, Applicati  2.89e-02
c    36     24     1.9     74 13  PCT-US95-1   Sequence 100, Applicat  2.89e-02
c    37     24     1.9     75 13  PCT-US95-1   Sequence 99, Applicati  2.89e-02
     38     24     1.9     81 13  PCT-US95-1   Sequence 92, Applicati  2.89e-02
c    39     24     1.9     91 13  PCT-US95-1   Sequence 95, Applicati  2.89e-02
     40     24     1.9    105  5  US-07-865-   Sequence 13, Applicati  2.89e-02
     41     24     1.9   2009  6  US-08-333-   Sequence 9, Applicatio  2.89e-02
     42     23     1.8     74 13  PCT-US95-1   Sequence 100, Applicat  1.06e-01
c    43     23     1.8    105  5  US-07-865-   Sequence 13, Applicati  1.06e-01
     44     23     1.8   1577 13  PCT-US95-1   Sequence 14, Applicati  1.06e-01
     45     23     1.8   1787 13  PCT-US95-1   Sequence 6, Applicatio  1.06e-01

                                 ALIGNMENTS

RESULT     1
ID   US-08-076-726-16 STANDARD; DNA; UNC, 4963 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 16, Application US/08076726.
CC   Sequence 16, Application US/08076726
CC   Patent No. 5464758
CC    GENERAL INFORMATION:
CC      APPLICANT:  Gossen, Manfred
CC      APPLICANT:  Bujard, Hermann
CC      TITLE OF INVENTION:  Tight Control of Gene Expression in
CC      TITLE OF INVENTION:  Eukaryotic Cells by Tetracycline-responsive Promot
  ers
CC      NUMBER OF SEQUENCES:  16
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  Sterne, Kessler, Goldstein & Fox
CC        STREET:  1100 New York Avenue, N.W.
CC        CITY:  Washington
CC        STATE:  District of Columbia
CC        COUNTRY:  United States of America
CC        ZIP:  20005-3934
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  US/08/076,726
CC        FILING DATE:  14-JUN-1993
CC        CLASSIFICATION:  435
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  Esmond, Robert W.
CC        REGISTRATION NUMBER:  32,893
CC        REFERENCE/DOCKET NUMBER:  0942.2490001
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  (202)371-2600
CC        TELEFAX:  (202)371-2540
CC        TELEX:  248636 SSK
CC    INFORMATION FOR SEQ ID NO:  16:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  4963 base pairs
CC        TYPE:  nucleic acid
CC        STRANDEDNESS:  both
CC         TOPOLOGY:  both
SQ   Sequence 4963 BP; 1259 A; 1300 C; 1278 G; 1126 T; 0 other;

  Query Match             10.8%;  Score 135;  DB 5;  Length 4963;
  Best Local Similarity   77.8%;  Pred. No. 2.62e-85;
  Matches  189;  Conservative    0;  Mismatches   54;  Indels    0;  Gaps    0;

Db   1005 GGAGACTCGCTACTGTGCAGTGTGCAATGACTATGCTTCAGGCTACCATTATGGAGTCTG 1064
          |||  ||| || ||| || || ||||  || || || || || || || |||||||||||
Qy    273 GGATGCTCACTTCTGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTG 332

Db   1065 GTCCTGTGAGGGCTGCAAGGCCTTCTTCAAGAGAAGTATTCAAGGACATAACGACTATAT 1124
          ||| ||||| || || |||||||| || || ||||| |||||||||||||| || |||||
Qy    333 GTCGTGTGAAGGATGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATAT 392

Db   1125 GTGTCCAGCCACCAACCAGTGCACCATTGATAAAAACAGGAGGAAGAGCTGCCAGGCCTG 1184
           |||||||| || || ||||| || || ||||||||| || | |||||||||||||||||
Qy    393 TTGTCCAGCTACAAATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTG 452

Db   1185 CCGGCTCCGCAAATGCTACGAAGTGGGAATGATGAAAGGTGGGATACGAAAAGACCGAAG 1244
          ||| || || || || ||||||||||||||| ||||  ||||    || | |||  || |
Qy    453 CCGACTTCGGAAGTGTTACGAAGTGGGAATGGTGAAGTGTGGCTCCCGGAGAGAGAGATG 512

Db   1245 AGG 1247
           ||
Qy    513 TGG 515


RESULT     2
ID   US-08-260-452-9 STANDARD; DNA; UNC; 4963 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 9, Application US/08260452.
CC   Sequence 9, Application US/08260452
CC   Patent No. 5650298
CC    GENERAL INFORMATION:
CC      APPLICANT:  Gossen, Manfred
CC      APPLICANT:  Bujard, Hermann
CC      APPLICANT:  Salfeld, Jochen
CC      APPLICANT:  Voss, Jeffrey
CC      TITLE OF INVENTION:  Tight Control of Gene Expression in Eucaryotic
CC      TITLE OF INVENTION:  Cells by Tetracycline-responsive Promoters
CC      NUMBER OF SEQUENCES:  10
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  Lahive & Cockfield
CC        STREET:  60 State Street
CC        CITY:  Boston
CC        STATE:  Massachusetts
CC        COUNTRY:  USA
CC        ZIP:  02109-1875
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  ASCII text
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  US/08/260,452
CC        FILING DATE:
CC        CLASSIFICATION:  435
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  08/076,327
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  Giulio A. DeConti, Jr.
CC        REGISTRATION NUMBER:  31,503
CC        REFERENCE/DOCKET NUMBER:  BBI-013CP
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  (617) 227-7400
CC        TELEFAX:  (617) 227-5941
CC    INFORMATION FOR SEQ ID NO:  9:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  4963 base pairs
CC        TYPE:  nucleic acid
CC        STRANDEDNESS:  double
```

```
CC         TOPOLOGY:  circular
CC      MOLECULE TYPE:  DNA (genomic)
CC      ORIGINAL SOURCE:
CC        ORGANISM:  Human cytomegalovirus
CC      IMMEDIATE SOURCE:
CC        CLONE:  pUHD BGR4
SQ   Sequence 4963 BP; 1259 A; 1300 C; 1278 G; 1126 T; 0 other;

  Query Match             10.8%;  Score 135;  DB 7;  Length 4963;
  Best Local Similarity   77.8%;  Pred. No. 2.62e-85;
  Matches  189;  Conservative    0;  Mismatches   54;  Indels    0;  Gaps    0;

Db   1005 GGAGACTCGCTACTGTGCAGTGTGCAATGACTATGCTTCAGGCTACCATTATGGAGTCTG 1064
          |||  ||| || ||| || || ||||  || || || || || || || |||||||||||
Qy    273 GGATGCTCACTTCTGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTG 332

Db   1065 GTCCTGTGAGGGCTGCAAGGCCTTCTTCAAGAGAAGTATTCAAGGACATAACGACTATAT 1124
          ||| ||||| || || |||||||| || || ||||| |||||||||||||| || |||||
Qy    333 GTCGTGTGAAGGATGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATAT 392

Db   1125 GTGTCCAGCCACCAACCAGTGCACCATTGATAAAAACAGGAGGAAGAGCTGCCAGGCCTG 1184
           |||||||| || || ||||| || || ||||||||| || | |||||||||||||||||
Qy    393 TTGTCCAGCTACAAATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTG 452

Db   1185 CCGGCTCCGCAAATGCTACGAAGTGGGAATGATGAAAGGTGGGATACGAAAAGACCGAAG 1244
          ||| || || || || ||||||||||||||| ||||  ||||    || | |||  || |
Qy    453 CCGACTTCGGAAGTGTTACGAAGTGGGAATGGTGAAGTGTGGCTCCCGGAGAGAGAGATG 512

Db   1245 AGG 1247
           ||
Qy    513 TGG 515


RESULT     3
ID   US-08-312-312A-1 STANDARD; DNA; UNC; 3460 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 1, Application US/08312312A
CC   Sequence 1, Application US/08312312A
CC   Patent No. 5550024
CC    GENERAL INFORMATION:
CC      APPLICANT:  Rothschild, Max F.
CC      APPLICANT:  Tuggle, Christopher K.
CC      APPLICANT:  Jacobson, Carol D.
CC      APPLICANT:  Mileham, Alan J.
CC      APPLICANT:  Plastow, Graham S.
CC      APPLICANT:  Vaske, David A.
CC      TITLE OF INVENTION:  Genetic Markers for Pig Litter Size
CC      NUMBER OF SEQUENCES:  9
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  Zarley, McKee, Thomte, Voorhees & Sease
CC        STREET:  801 Grand Suite 3200
CC        CITY:  Des Moines
CC        STATE:  Iowa
CC        COUNTRY:  United States
CC        ZIP:  50309
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  PatentIn Release #1.0, Version #1.30
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  US/08/312,312A
CC        FILING DATE:  26-SEP-1994
CC        CLASSIFICATION:  435
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US 07/961,819
CC        FILING DATE:  19-OCT-1992
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US 07/687,708
CC        FILING DATE:  19-APR-1991
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  Nebel, Heidi S.
CC        REGISTRATION NUMBER: 37,719
CC        REFERENCE/DOCKET NUMBER: BPDC-4
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE: 515-288-3667
CC        TELEFAX: 515-288-1338
CC    INFORMATION FOR SEQ ID NO: 1:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH: 3460 base pairs
CC        TYPE: nucleic acid
CC        STRANDEDNESS: double
CC        TOPOLOGY: linear
CC      MOLECULE TYPE: DNA (genomic)
CC      HYPOTHETICAL: NO
CC      ANTI-SENSE: NO
CC      FEATURE:
CC        NAME/KEY: misc_feature
CC        LOCATION: 303..304
CC        OTHER INFORMATION: /note= "Gap in DNA sequence of
CC        OTHER INFORMATION: undetermined length"
CC      FEATURE:
CC        NAME/KEY: misc_feature
CC        LOCATION: 1197..1198
CC        OTHER INFORMATION: /note= "Gap in DNA sequence of
CC        OTHER INFORMATION: undetermined length"
CC      FEATURE:
CC        NAME/KEY: exon
CC        LOCATION: 534..600
SQ   Sequence 3460 BP; 1052 A; 765 C; 634 G; 1009 T; 0 other;

  Query Match             5.7%;  Score 71;  DB 6;  Length 3460;
  Best Local Similarity   83.3%;  Pred. No. 5.58e-35;
  Matches   95;  Conservative    0;  Mismatches  18;  Indels   1;  Gaps   1;

Db     490 TTCATCATGCCCACTTCGTAGCACTTGGGTAGCCGGCAGGC-TGACAGCTCTTCCTCCTG 548
           |||| ||| |||||||||||| ||||| || || ||||||||| || |||||||| | || |
Qy     488 TTCACCATTCCCACTTCGTAACACTTCCGAAGTCGGCAGGCCTGGCAGCTCTTGCGCCGG 429

Db     549 TTCTTATCAATTGTGCACTGGTTGGTAGCTGGACACATGTAGTCATTATGTCCT 602
           || ||||| ||||| ||||| || ||||||||||| || || ||||||||||||
Qy     428 TTTTTATCGATTGTACACTGATTTGTAGCTGGACAAATATAATCATTATGTCCT 375


RESULT    4
ID   US-08-076-726-15 STANDARD; DNA; UNC; 6244 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 15, Application US/08076726.
CC   Sequence 15, Application US/08076726
CC   Patent No. 5464758
CC    GENERAL INFORMATION:
CC      APPLICANT: Gossen, Manfred
CC      APPLICANT: Bujard, Hermann
CC      TITLE OF INVENTION: Tight Control of Gene Expression in
CC      TITLE OF INVENTION: Eucaryotic Cells by Tetracycline-responsive Promot
  ers
CC      NUMBER OF SEQUENCES: 16
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE: Sterne, Kessler, Goldstein & Fox
CC        STREET: 1100 New York Avenue, N.W
CC        CITY: Washington
CC        STATE: District of Columbia
CC        COUNTRY: United States of America
CC        ZIP: 20005-3934
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE: Floppy disk
CC        COMPUTER: IBM PC compatible
CC        OPERATING SYSTEM: PC-DOS/MS-DOS
CC        SOFTWARE: PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER: US/08/076,726
CC        FILING DATE: 14-JUN-1993
CC        CLASSIFICATION: 435
CC      ATTORNEY/AGENT INFORMATION:
```

```
CC        NAME: Esmond, Robert W.
CC        REGISTRATION NUMBER: 32,893
CC        REFERENCE/DOCKET NUMBER: 0942.2490001
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE: (202)371-2600
CC        TELEFAX: (202)371-2540
CC        TELEX: 248636 SSK
CC    INFORMATION FOR SEQ ID NO: 15:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH: 6244 base pairs
CC        TYPE: nucleic acid
CC        STRANDEDNESS: both
CC        TOPOLOGY: both
SQ   Sequence 6244 BP; 1493 A; 1749 C; 1610 G; 1402 T; 0 other;

  Query Match             5.0%;  Score 63;  DB 5;  Length 6244;
  Best Local Similarity   66.3%;  Pred. No. 5.80e-29;
  Matches  128;  Conservative    0;  Mismatches  65;  Indels   0;  Gaps   0;

Db    2365 TCTGTGGGATGAAGCATCAGGCTGTCATTATGGTGTCCTCACCTGTGGGAGCTGTAAGG 2424
           ||||  | ||| | ||||| || | ||| ||||| |||     | ||||    | |||||||
Qy     293 TCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATGTAAGG 352

Db    2425 TCTTCTTTAAAAGGGCAATGGAAGGGCAGCATAACTATTTATCTGCTGGAAGAAATGACT 2484
            ||| ||||||||     ||   |||| ||   || | ||| | ||| | |   | |||| | |
Qy     353 CCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCAGT 412

Db    2485 GCATTGTTGATAAAATCCGCAGGAAAAACTGCCCGGCGTGTCGGTTTAGAAATTGCTGTC 2544
           | |   | ||||||| |||   | || | ||||| ||| || || ||| | ||||| |
Qy     413 GTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTACG 472

Db    2545 AACCTGGCATGGT 2557
           |||  || |||||
Qy     473 AAGTGGGAATGGT 485


RESULT    5
ID   US-08-260-452-8 STANDARD; DNA; UNC; 6244 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 8, Application US/08260452
CC   Sequence 8, Application US/08260452
CC   Patent No. 5650298
CC    GENERAL INFORMATION:
CC      APPLICANT: Gossen, Manfred
CC      APPLICANT: Bujard, Hermann
CC      APPLICANT: Salfeld, Jochen
CC      APPLICANT: Voss, Jeffrey
CC      TITLE OF INVENTION: Tight Control of Gene Expression in Eucaryotic
CC      TITLE OF INVENTION: Cells by Tetracycline-responsive Promoters
CC      NUMBER OF SEQUENCES: 10
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE: Lahive & Cockfield
CC        STREET: 60 State Street
CC        CITY: Boston
CC        STATE: Massachusetts
CC        COUNTRY: USA
CC        ZIP: 02109-1875
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE: Floppy disk
CC        COMPUTER: IBM PC compatible
CC        OPERATING SYSTEM: PC-DOS/MS-DOS
CC        SOFTWARE: ASCII text
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER: US/08/260,452
CC        FILING DATE:
CC        CLASSIFICATION: 435
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER: 08/076,327
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME: Giulio A. DeConti, Jr.
CC        REGISTRATION NUMBER: 31,503
CC        REFERENCE/DOCKET NUMBER: BBI-013CP
CC        TELECOMMUNICATION INFORMATION:
CC          TELEPHONE:  (617) 227-7400
CC          TELEFAX:  (617) 227-5941
CC      INFORMATION FOR SEQ ID NO:  8:
CC        SEQUENCE CHARACTERISTICS:
CC          LENGTH:  6244 base pairs
CC          TYPE:  nucleic acid
CC          STRANDEDNESS:  double
CC          TOPOLOGY:  circular
CC        MOLECULE TYPE:  DNA (genomic)
CC        ORIGINAL SOURCE:
CC          ORGANISM:  Human cytomegalovirus
CC          STRAIN:  Towne (hCMV)
CC        IMMEDIATE SOURCE:
CC          CLONE:  pUHD BGR3
SQ   Sequence 6244 BP; 1483 A; 1749 C; 1610 G; 1402 T; 0 other;

  Query Match             5.0%;  Score 63;  DB 7;  Length 6244;
  Best Local Similarity  66.3%;  Pred. No. 5.80e-29;
  Matches  128;  Conservative    0;  Mismatches  65;  Indels    0;  Gaps    0;

Db    2365 TCTGTGGGGATGAAGCATCAGGCTGTCATTATGGTGTCCTCACCTGTGGGAGCTGTAAGG 2424
           ||||  | ||| | ||||| || | ||| ||||| |||      | ||||    | |||||||
Qy     293 TCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATGTAAGG 352

Db    2425 TCTTCTTTAAAAGGGCAATGGAAGGGCAGCATAACTATTTATGTGCTGGAAGAAATGACT 2484
             ||| ||||||||      ||   |||| ||   || | ||| | ||| | |   | ||||| | |
Qy     353 CCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCAGT 412

Db    2485 GCATTGTTGATAAAATGGGCAGGAAAAACTGGCCGGCGGTGTCCCTTAGAAACTGCTGTC 2544
           | |    | ||||||| |||   | || | ||||| ||| || || ||| | ||||| |
Qy     413 GTACAATCGATAAAAACGGGAGGCAAGAGCTGGCAGGCCTGCCGATTCGGAAGTGTTACG 472

Db    2545 AAGCTGGCATGGT 2557
           |||  || |||||
Qy     473 AAGTGGGAATGGT 485


RESULT     6
ID   US-08-232 463 14 STANDARD; DNA; UNC; 7218 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 14, Application US/08232463.
CC   Sequence 14, Application US/08232463
CC   Patent No. 5670367
CC    GENERAL INFORMATION:
CC      APPLICANT:  DORNER, F.
CC      APPLICANT:  SCHEIFLINGER, F
CC      APPLICANT:  FALKNER, F. G.
CC      TITLE OF INVENTION:  RECOMBINANT FOWLPOX VIRUS
CC      NUMBER OF SEQUENCES:  52
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  Foley & Lardner
CC        STREET:  1800 Diagonal Road, Suite 500
CC        CITY:  Alexandria
CC        STATE:  VA
CC        COUNTRY:  USA
CC        ZIP:  22313-0299
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  US/08/232,463
CC        FILING DATE:
CC        CLASSIFICATION:  435
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US/07/935,313
CC        FILING DATE:
CC        APPLICATION NUMBER:  EP 91 114 300.6
CC        FILING DATE:  25 AUG 1991
CC      ATTORNEY/AGENT INFORMATION:
```

```
CC        NAME:  BENT, Stephen A.
CC        REGISTRATION NUMBER:  29,768
CC        REFERENCE/DOCKET NUMBER:  30472/114 IMMU
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  (703)836-9300
CC        TELEFAX:  (703)683-4109
CC        TELEX:  899149
CC    INFORMATION FOR SEQ ID NO:  14:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  7218 base pairs
CC        TYPE:  nucleic acid
CC        STRANDEDNESS:  single
CC        TOPOLOGY:  linear
CC      IMMEDIATE SOURCE:
CC        CLONE:  pT2gpt-F1s
SQ   Sequence 7218 BP; 1944 A; 1491 C; 1486 G; 1929 T; 368 other;

  Query Match             5.0%;  Score 62;  DB 7;  Length 7218;
  Best Local Similarity   1.4%;  Pred. No. 3.22e-28;
  Matches    5;  Conservative  205;  Mismatches 148;  Indels    0;  Gaps    0;

Db    1083 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1142
           ::::: : :::      : : :   :: :: ::::::::   :::: ::   : :   : :  ::
Cp     612 CTCCCGCACTCGGGGGCGGTCGGCCCCACTTCTCTTCGCCTTGCCGGCACAGTGCAGCTG 553

Db    1143 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1202
           ::: ::  : ::::::: ::::: : :      :   :  ::: : :::::::::::      ::
Cp     552 CTCGTCGGCACTTCTCTGTCTCCGCACAAGGCGGTACCCACATCTCTCTCTCCGGGAGCC 493

Db    1203 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1262
             : :::: :: ::::: :::: :   : :::::      ::   :    :::   :   ::::: :
Cp     492 ACACTTCACCATTCCCACTTCGTAACACTTCCGAAGTCGGCACGCCTGCCAGCTCTTGCG 433

Db    1263 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1322
           ::   ::::: ::   :: : : ::   ::: :   ::     :     : :   :: :: : :::::
Cp     432 CCGGTTTTTATCGATTGTACACTGATTTGTAGCTGGACAAATATAATCATTATGTCCTTG 373

Db    1323 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1382
             : ::::::::          :::: : :::::: : :   ::    :::: :    :    : :::
Cp     372 AATGCTTCTTTTAAAAAAGGCCTTACATCCTTCACACGACCAGACTCCATAGTGATATCC 313

Db    1383 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYGTACCA 1440
           .  .  .  .   . . .: :    :   : :      :    : :::::::::    :::| ||||
Cp     312 CGATGCGTAATCGCTGCAGACAGCGCAGAAGTGAGCATCCCTCTTTGAACCTGGACCA 255


RESULT     7
ID   5223606-1 STANDARD; DNA; UNC; 3238 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Patent No. 5223606.
CC   Patent No. 5223606
CC      APPLICANT:  BLAUDIN DE THE, HUGHES;MARCHIO, AGNES;TIOLLAIS,
CC   PIERRE;DEJEAN, ANNE
CC      TITLE OF INVENTION: STEROID/THYROID HORMONE RECEPTOR-RELATED
CC   PROTEIN INAPPROPRIATELY EXPRESSED IN HUMAN HEPATOCELLULAR CARCINOMA
CC      NUMBER OF SEQUENCES: 11
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  US/07/134,130
CC        FILING DATE: 17-DEC-1987
CC      PRIOR APPLICATION DATA:
CC   SEQ ID NO:1:
CC        LENGTH: 2989
SQ   Sequence 3238 BP; 300 A; 648 C; 600 G; 841 T; 249 other;

  Query Match             4.4%;  Score 55;  DB 1;  Length 2989;
  Best Local Similarity  65.0%;  Pred. No. 4.64e-23;
  Matches  130;  Conservative    0;  Mismatches  69;  Indels    1;  Gaps    1;

Db     560 CTGCTTCGTCTGCCAGGACAAATCATCAGGGTACCACTATGGGGTCAGCGCCTGTGAGGG 619
           ||||    ||||||    ||   |    |||| || || |||||||| ||| |   | ||||| ||
Qy     285 CTGGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGG 344
```

```
Db    620 ATGTAAGGGCTTTTTCGGCAGAAGTATTCAGAAGAATATGATTTACACTTGTCACCGAGA 679
          |||||||| ||||||    ||||| |||||     |||    ||| | |||||
Qy    345 ATGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTAC 404

Db    680 TAAGAACTGTGTTATTAATAAAAGTCACCAGGAATCGATGCCAATACTGTCGACTCCAGA 739
           ||  | |||   ||  ||||||  |  | | ||  | |||||   |||  ||||| | ||
Qy    405 AAATCAGTGTACAATCGATAAAAACCGGC-GCAAGAGCTGCCAGGCCTGCCGACTTCGGA 463

Db    740 AGTGCTTTGAAGTGGGAATG 759
          |||| |  ||||||||||||
Qy    464 AGTGTTACGAAGTGGGAATG 483

RESULT    8
ID    US-08-629-939-1 STANDARD; DNA; UNC; 3014 BP
AC    xxxxxx
DT    01-JAN-1900
DE    Sequence 1, Application US/08629939
CC    Sequence 1, Application US/08629939
CC    Patent No. 5645995
CC     GENERAL INFORMATION:
CC       APPLICANT:  Kieback, Dirk G.
CC       TITLE OF INVENTION:  METHODS FOR DIAGNOSING AN INCREASED
CC       TITLE OF INVENTION:  RISK OF BREAST OR OVARIAN CANCER
CC       NUMBER OF SEQUENCES:  14
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  SUGHRUE, MION, ZINN, MACPEAK & SEAS
CC         STREET:  2100 Pennsylvania Avenue, N.W., Suite 800
CC         CITY:  Washington, D.C.
CC         STATE:  D.C.
CC         COUNTRY:  U.S.A.
CC         ZIP:  20037
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/629,939
CC         FILING DATE:  12-APRIL-1996
CC         CLASSIFICATION:  435
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  KIT, Gordon
CC         REGISTRATION NUMBER:  30,764
CC         REFERENCE/DOCKET NUMBER:  A-6612
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (202) 293-7060
CC         TELEFAX:  (202) 293-7860
CC     INFORMATION FOR SEQ ID NO:  1:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  3014 base pairs
CC         TYPE:  nucleic acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  cDNA
CC       HYPOTHETICAL:  NO
SQ    Sequence 3014 BP; 623 A; 929 C; 837 G; 625 T; 0 other.

  Query Match              4.2%;  Score 53;  DB 7;  Length 3014;
  Best Local Similarity  63.7%;  Pred. No. 1.32e-21;
  Matches  123;  Conservative    0;  Mismatches  70;  Indels    0;  Gaps    0;

Db   1881 TCTGTGGGGATGAAGCATCAGGCTGTCATTATGCTGTCCTTACCTGTGGGAGCTGTAAGG 1940
          ||||  | ||| | | ||  || | ||| |||  |||    | ||||   | |||||||
Qy    293 TCTGCAGTGATTACGGATTGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATGTAAGG 352

Db   1941 TCTTCTTTAAGAGGGCAATGGAAGGGCAGCACAACTACTTATGTGCTGGAAGAAATGACT 2000
           ||| ||||| ||    ||  |||| ||  |  | ||  | ||| | |  | |||| | |
Qy    353 CCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCAGT 412

Db   2001 GCATCGTTGATAAAATCCGCAGAAAAAACTGCCCAGCATGTCGGCTTAGAAAGTGCTGTC 2060
          | |   | |||||||  |||  | || | |||||  || || || ||| | ||||| |
Qy    413 GTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTACG 472
```

```
Db   2061 AGGCTGGCATGGT 2073
          | |  || |||||
Qy    473 AAGTGGGAATGGT 485

RESULT    9
ID    5260432-1 STANDARD; DNA; UNC; 1707 BP.
AC    xxxxxx
DT    01-JAN-1900
DE    Patent No. 5260432.
CC    Patent No. 5260432
CC       APPLICANT: TAKAKU, FUMIMARO;ISHIKAWA, TAKASHI;IMAWARI, MICHIO;
CC    EVANS, RONALD M.;UMESONO, KAZUHIKO
CC       TITLE OF INVENTION: HUMAN GAMMA RETINOIC ACID RECEPTOR DNA
CC       NUMBER OF SEQUENCES: 4
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER: US/07/370,407
CC         FILING DATE: 22-JUN-1989
CC    SEQ ID NO:1:
CC         LENGTH: 1576
SQ    Sequence 1707 BP; 346 A; 478 C; 467 G; 285 T; 131 other;

  Query Match              4.1%;  Score 51;  DB 1;  Length 1576;
  Best Local Similarity  65.8%;  Pred. No. 3.68e-20;
  Matches  127;  Conservative    0;  Mismatches  64;  Indels    2;  Gaps    2;

Db    473 GTGTGCAATGACAAGTCCTCTGGCTACCACTATGGGGTCAGCTCTTGTGAAGGCTGCAAG 532
          || |||| |||  |    |  || || |||||||| ||| | || |||||||| || |||
Qy    292 GTCTGCAGTGATTACGGATTGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATGTAAG 351

Db    533 GGCTTCTTTCGGCGAAGCATCCAGAAGAACATGGTGTACACGTGTCACCGCGACAAAA-A 591
          | ||| |||    ||||||| ||     | |  |  || |  |||| | || |||||  |
Qy    352 GCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTC-CAGCTACAAATCA 410

Db    592 CTGTATCATCAACAAGGTGACCAGGAATCGCTGCCAGTACTGCCGGCTACAGAAGTGCTT 651
           ||||  ||| | ||        | ||  ||||||||  |||||| || | |||||| |
Qy    411 GTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTA 470

Db    652 CGAAGTGGGCATG 664
          ||||||||| |||
Qy    471 CGAAGTGGGAATG 483

RESULT   10
ID    US-08-087-151-1 STANDARD; DNA; UNC; 387 BP
AC    xxxxxx
DT    01-JAN-1900
DE    Sequence 1, Application US/08087151.
CC    Sequence 1, Application US/08087151
CC    Patent No. 5468624
CC     GENERAL INFORMATION:
CC       APPLICANT:  E. Brad Thompson
CC       APPLICANT:  Lynne V. Nazareth
CC       TITLE OF INVENTION:  ANTI-TUMOR ACTIVITY OF A MODIFIED FRAGMENT
CC       TITLE OF INVENTION:  OF THE GLUCOCORTICOID RECEPTOR
CC       NUMBER OF SEQUENCES:  3
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Arnold, White & Durkee
CC         STREET:  P.O. Box 4433
CC         CITY:  Houston
CC         STATE:  Texas
CC         COUNTRY:  USA
CC         ZIP:  77210
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  WordPerfect 5.1
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/087,151
CC         FILING DATE:  July 7, 1993
CC         CLASSIFICATION:  435
CC        ATTORNEY/AGENT INFORMATION:
CC          NAME:  Hodgins, Daniel S.
CC          REGISTRATION NUMBER:  31,026
CC          REFERENCE/DOCKET NUMBER:  UTSC-141/HOD
CC        TELECOMMUNICATION INFORMATION:
CC          TELEPHONE:  512-320-7200
CC          TELEFAX:  512-474-7577
CC          TELEX:  No. 5468624 Applicable
CC      INFORMATION FOR SEQ ID NO:  1:
CC        SEQUENCE CHARACTERISTICS:
CC          LENGTH:  387 base pairs
CC          TYPE:  nucleic acid
CC          STRANDEDNESS:  double
CC          TOPOLOGY:  linear
CC        MOLECULE TYPE:  genomic DNA
SQ   Sequence 387 BP; 109 A; 92 C; 88 G; 98 T; 0 other;

  Query Match              3.8%;  Score 47;  DB 5;  Length 387;
  Best Local Similarity  63.0%;  Pred. No. 2.64e-17;
  Matches  126;  Conservative    0;  Mismatches  73;  Indels    1;  Gaps    1;

Db     188 TCTGCCTGGTGTGCTCTGATGAAGCTTCAGGATGTCATTATGGAGTCTTAACTTGTGGAA 247
           |||||   || |||   ||| | || || |||| ||| ||||||||||   | |||| |
Qy     284 TCTGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAG 343

Db     248 GCTGTAAAGTTTTCTTCAAAAGAGCAGTGGAAGGACAGCACAATTACCTATGTGCTGGAA 307
           | ||||| |  || || ||||||    |  ||||||| |  ||||  | ||| | |  |
Qy     344 GATGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTA 403

Db     308 GGAATGATTGCATCATCGCGATAAAATT-CGAAAAAACTGCCCAGCATGCCGCTATCGAA 366
             ||| | || |  ||||  | |||    || || | |||||  || |||||   ||| |
Qy     404 CAAATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGA 463

Db     367 AATGTCTTCAGGCTGGAATG 386
           | |||    | |  ||||||
Qy     464 AGTGTTACGAAGTGGGAATG 483


RESULT   11
ID   US-08-561-072-1 STANDARD; DNA; UNC; 387 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 1, Application US/08561072
CC   Sequence 1, Application US/08561072
CC   Patent No. 5571791
CC    GENERAL INFORMATION:
CC      APPLICANT:  E. Brad Thompson
CC      APPLICANT:  Lynne V. Nazareth
CC      TITLE OF INVENTION:  ANTI-TUMOR ACTIVITY OF A MODIFIED FRAGMENT
CC      TITLE OF INVENTION:  OF THE GLUCOCORTICOID RECEPTOR
CC      NUMBER OF SEQUENCES:  3
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  Arnold, White & Durkee
CC        STREET:  P.O. Box 4433
CC        CITY:  Houston
CC        STATE:  Texas
CC        COUNTRY:  USA
CC        ZIP:  77210
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  WordPerfect 5.1
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  US/08/561,072
CC        FILING DATE:  20-NOV-1995
CC        CLASSIFICATION:  530
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  08/087,151
CC        FILING DATE:  July 7, 1993
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  Hodgins, Daniel S.
CC        REGISTRATION NUMBER:  31,026
CC          REFERENCE/DOCKET NUMBER:  UTSC-141/HOD
CC        TELECOMMUNICATION INFORMATION:
CC          TELEPHONE:  512-320-7200
CC          TELEFAX:  512-474-7577
CC          TELEX:  No. 5571791 Applicable
CC      INFORMATION FOR SEQ ID NO:  1:
CC        SEQUENCE CHARACTERISTICS:
CC          LENGTH:  387 base pairs
CC          TYPE:  nucleic acid
CC          STRANDEDNESS:  double
CC          TOPOLOGY:  linear
CC        MOLECULE TYPE:  genomic DNA
SQ   Sequence 387 BP; 109 A; 92 C; 88 G; 98 T; 0 other;

  Query Match              3.8%;  Score 47;  DB 6;  Length 387;
  Best Local Similarity  63.0%;  Pred. No. 2.64e-17;
  Matches  126;  Conservative    0;  Mismatches  73;  Indels    1;  Gaps    1;

Db     188 TCTGCCTGGTGTGCTCTGATGAAGCTTCAGGATGTCATTATGGAGTCTTAACTTGTGGAA 247
           |||||   || |||   ||| | || || |||| ||| ||||||||||   | |||| |
Qy     284 TCTGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAG 343

Db     248 GCTGTAAAGTTTTCTTCAAAAGAGCAGTGGAAGGACAGCACAATTACCTATGTGCTGGAA 307
           | ||||| |  || || ||||||    |  ||||||| |  ||||  | ||| | |  |
Qy     344 GATGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTA 403

Db     308 GGAATGATTGCATCATCGCGATAAAATT-CGAAAAAACTGCCCAGCATGCCGCTATCGAA 366
             ||| | || |  ||||  | |||    || || | |||||  || |||||   ||| |
Qy     404 CAAATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGA 463

Db     367 AATGTCTTCAGGCTGGAATG 386
           | |||    | |  ||||||
Qy     464 AGTGTTACGAAGTGGGAATG 483


RESULT   12
ID   PCT-US92-02320A-3 STANDARD; DNA; UNC; 2928 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 3, Application PC/TUS9202320A.
CC   Sequence 3, Application PC/TUS9202320A
CC    GENERAL INFORMATION:
CC      APPLICANT:  Sloan-Kettering Institute, For Cancer Research
CC      TITLE OF INVENTION.  METHODS FOR DETECTION AND TREATMENT OF CANCER
CC      NUMBER OF SEQUENCES:  6
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  COOPER & DUNHAM
CC        STREET:  30 ROCKEFELLER PLAZA
CC        CITY:  NEW YORK
CC        STATE:  NEW YORK
CC        COUNTRY:  U.S.A.
CC        ZIP:  10112
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  PatentIn Release #1.24
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  PCT/US92/02320A
CC        FILING DATE:
CC        CLASSIFICATION:
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US 675,084
CC        FILING DATE:  22-MAR-1991
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US 673,838
CC        FILING DATE:  22-MAR-1991
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  WHITE, JOHN P
CC        REGISTRATION NUMBER:  28,678
CC        REFERENCE/DOCKET NUMBER:  38694-PCT
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  (212) 977-9550
CC          TELEFAX:  (212) 644-0525
CC          TELEX:  (212) 422523 COOP UI
CC      INFORMATION FOR SEQ ID NO:  3:
CC        SEQUENCE CHARACTERISTICS:
CC          LENGTH:  2928 base pairs
CC          TYPE:  nucleic acid
CC          STRANDEDNESS:  single
CC          TOPOLOGY:  linear
CC        MOLECULE TYPE:  cDNA
CC        IMMEDIATE SOURCE:
CC          CLONE:  hRAR ALPHA
CC        FEATURE:
CC          NAME/KEY:  CDS
CC          LOCATION:  103..1499
CC          OTHER INFORMATION:
SQ   Sequence 2928 BP; 582 A; 1011 C; 789 G; 546 T; 0 other;

  Query Match             3.4%;  Score 42;  DB 10;  Length 2928;
  Best Local Similarity   60.6%;  Pred. No. 8.32e-14;
  Matches  120;  Conservative    0;  Mismatches  78;  Indels    0;  Gaps    0;

Db    364 TGCTTTGTCTGTCAGGACAAGTCCTCAGGCTACTACTATGGGGTCAGCGCCTGTGAGGGC 423
          |||  ||||||    ||  |  | || || || |||||||| ||| |  | ||||| ||
Qy    286 TGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGA 345

Db    424 TGCAAGGGCTTCTTCCGCCGCAGCATCCAGAAGAACATGGTGTACACGTGTCACCGGGAC 483
          || |||| ||| ||     | ||||| ||     | |  |  || |  ||||
Qy    346 TGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACA 405

Db    484 AAGAACTGCATCATCAACAAGGTGACCCGGAACCGCTGCCAGTACTGCCGACTGCAGAAG 543
          ||  | || |  ||| | ||       || ||  ||||||||  ||||||||| | ||||
Qy    406 AATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAG 465

Db    544 TGCTTTGAAGTGGGCATG 561
          || |  |||||||| |||
Qy    466 TGTTACGAAGTGGGAATG 483


RESULT   13
ID   5171671-1 STANDARD; DNA; UNC; 3185 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Patent No. 5171671.
CC  Patent No. 5171671
CC        APPLICANT: EVANS, RONALD M.;ONG, ESTELITA S.;SEGUI,
CC  PRUDIMAR S.;THOMPSON, CATHERINE C.;UEMSONO, KAZUHIKO
CC  GUGUERE, VINCENT
CC        TITLE OF INVENTION: RETINOIC ACID RECEPTOR COMPOSITION
CC        NUMBER OF SEQUENCES: 2
CC        CURRENT APPLICATION DATA:
CC          APPLICATION NUMBER:  US/07/546,256
CC          FILING DATE: 06-AUG-1990
CC        PRIOR APPLICATION DATA:
CC          APPLICATION NUMBER: 276,536
CC          FILING DATE: 30-NOV-1988
CC          APPLICATION NUMBER: 128,331
CC          FILING DATE: 02-DEC-1987
CC  SEQ ID NO:1:
CC          LENGTH: 2940
SQ   Sequence 3185 BP; 594 A; 1010 C; 790 G; 546 T; 245 other;

  Query Match             3.4%;  Score 42;  DB 1;  Length 2940;
  Best Local Similarity   60.6%;  Pred. No. 8.32e-14;
  Matches  120;  Conservative    0;  Mismatches  78;  Indels    0;  Gaps    0;

Db    364 TGCTTTGTCTGTCAGGACAAGTCCTCAGGCTACTACTATGGGGTCAGCGCCTGTGAGGGC 423
          |||  ||||||    ||  |  | || || || |||||||| ||| |  | ||||| ||
Qy    286 TGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGA 345

Db    424 TGCAAGGGCTTCTTCCGCCGCAGCATCCAGAAGAACATGGTGTACACGTGTCACCGGGAC 483
          || |||| ||| ||     | ||||| ||     | |  |  || |  ||||
Qy    346 TGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACA 405

Db    484 AAGAACTGCATCATCAACAAGGTGACCCGGAACCGCTGCCAGTACTGCCGACTGCAGAAG 543
          ||  | || |  ||| | ||       || ||  ||||||||  ||||||||| | ||||
Qy    406 AATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAG 465

Db    544 TGCTTTGAAGTGGGCATG 561
          || |  |||||||| |||
Qy    466 TGTTACGAAGTGGGAATG 483


RESULT   14
ID   PCT-US92-02320A-1 STANDARD; DNA; UNC; 3036 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 1, Application PC/TUS9202320A.
CC   Sequence 1, Application PC/TUS9202320A
CC    GENERAL INFORMATION:
CC      APPLICANT:  Sloan-Kettering Institute, For Cancer Research
CC      TITLE OF INVENTION:  METHODS FOR DETECTION AND TREATMENT OF CANCER
CC      NUMBER OF SEQUENCES:  6
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  COOPER & DUNHAM
CC        STREET:  30 ROCKEFELLER PLAZA
CC        CITY:  NEW YORK
CC        STATE:  NEW YORK
CC        COUNTRY:  U.S.A.
CC        ZIP:  10112
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  PatentIn Release #1.24
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  PCT/US92/02320A
CC        FILING DATE:
CC        CLASSIFICATION:
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US 675,084
CC        FILING DATE:  22-MAR-1991
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US 673,838
CC        FILING DATE:  22-MAR-1991
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  WHITE, JOHN P
CC        REGISTRATION NUMBER:  28,678
CC        REFERENCE/DOCKET NUMBER:  38694-PCT
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  (212) 977-9550
CC        TELEFAX:  (212) 644-0525
CC        TELEX:  (212) 422523 COOP UI
CC    INFORMATION FOR SEQ ID NO:  1:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  3036 base pairs
CC        TYPE:  nucleic acid
CC        STRANDEDNESS:  single
CC        TOPOLOGY:  linear
CC      MOLECULE TYPE:  cDNA
CC      IMMEDIATE SOURCE:
CC        CLONE:  MYL-RAR
CC      FEATURE:
CC        NAME/KEY:  CDS
CC        LOCATION:  67..2457
CC        OTHER INFORMATION:
SQ   Sequence 3036 BP; [illegible] A; 1116 C; [illegible] G; 523 T; 0 other;

  Query Match             3.4%;  Score 42;  DB 10;  Length 3036;
  Best Local Similarity   60.6%;  Pred. No. 8.32e-14;
  Matches  120;  Conservative    0;  Mismatches  78;  Indels    0;  Gaps    0;

Db   1333 TGCTTTGTCTGTCAGGACAAGTCCTCAGGCTACTACTATGGGGTCAGCGCCTGTGAGGGC 1392
          |||  ||||||    ||  |  | || || || |||||||| ||| |  | ||||| ||
Qy    286 TGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGA 345

Db   1393 TGCAAGGGCTTCTTCCGCCGCAGCATCCAGAAGAACATGGTGTACACGTGTCACCGGGAC 1452
```

```
           || |||| ||| |      | ||||| ||     | |  |  || |  ||||
Qy     346 TGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACA 405

Db    1453 AAGAACTGCATCATCAACAAGGTGACCCGGAACCGCTGCCAGTACTGCCGACTGCAGAAG 1512
           ||  | || |  ||| | ||       || ||  ||||||||  ||||||||| | ||||
Qy     406 AATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAG 465

Db    1513 TGCTTTGAAGTGGGCATG 1530
           || |  |||||||| |||
Qy     466 TGTTACGAAGTGGGAATG 483


RESULT   15
ID   PCT-US91-00394-5 STANDARD, DNA, UNC, 1662 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 5, Application PC/TUS9100399.
CC   Sequence 5, Application PC/TUS9100399
CC    GENERAL INFORMATION:
CC      APPLICANT:  Mangelsdorf, Dr., David J.
CC      APPLICANT:  Evans Dr., Ronald M.
CC      TITLE OF INVENTION:  RETINOID RECEPTOR COMPOSITIONS AND
CC      TITLE OF INVENTION:  METHODS
CC      NUMBER OF SEQUENCES:  7
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  FITCH, EVEN, TABIN & FLANNERY
CC        STREET:  135 South LaSalle Street, Suite 900
CC        CITY:  Chicago
CC        STATE:  Illinois
CC        COUNTRY:  USA
CC        ZIP:  60603
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  PCT/US91/00399
CC        FILING DATE:  19910122
CC        CLASSIFICATION:  435
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  07/478,071
CC        FILING DATE:  09-FEB-1990
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  Reiter, Mr., Stephen E.
CC        REGISTRATION NUMBER:  31192
CC        REFERENCE/DOCKET NUMBER:  50852
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  (619) 552-1311
CC        TELEFAX:  (619) 552-0095
CC        TELEX:  20 6566 PATLAW CGO
CC    INFORMATION FOR SEQ ID NO:  5:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  1662 base pairs
CC        TYPE:  NUCLEIC ACID
CC        STRANDEDNESS:  single
CC        TOPOLOGY:  linear
CC      MOLECULE TYPE:  DNA (genomic)
CC      IMMEDIATE SOURCE:
CC        CLONE:  RXR MOUSE GAMMA
CC      FEATURE:
CC        NAME/KEY:  CDS
CC        LOCATION:  join(123..1514)
SQ   Sequence 1662 BP; 397 A; 483 C; 441 G; 341 T; 0 other;

  Query Match              3.0%;  Score 38;  DB 9;  Length 1662;
  Best Local Similarity  60.1%;  Pred. No. 4.46e-11;
  Matches   113;  Conservative    0;  Mismatches  75;  Indels    0;  Gaps    0;

Db     531 CACATCTGTGCCATCTGTGGGGACAGATCCTCAGGGAAGCACTACGGTGTGTACAGCTGT 590
           ||| |||| ||  ||||  |  |     | || ||  | ||||| || || |     |||
Qy     280 CACTTCTGCGCTGTCTGCAGCCATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGT 339
```

```
Db     591 GAAGGTTGCAAAGGCTTCTTCAAAAGGACCATCAGGAAAGATCTCATCTACACCTGTCGG 650
           ||||| || || | ||| || ||||| | |||      | ||      || |  ||||
Qy     340 GAAGGATGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCA 399

Db     651 GATAACAAAGATTGTCTCATCGACAAGCGCCAGCGCAACCGCTGCCAGTACTGTCGCTAC 710
           | ||  ||  | |||   ||||| ||   || ||||||  ||||||||  ||| ||
Qy     400 GCTACAAATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTT 459

Db     711 CAGAAGTG 718
           | ||||||
Qy     460 CGGAAGTG 467

Search completed. Thu Dec 18 19.16.29 1997
Job time : 66 secs.
```

```
*****************************************************************
  __  __   ____    ____   ____    ____   _   _
 |  \/  | |  _ \  / ___| |  _ \  / ___| | | | |
 | |\/| | | |_) | \___ \ | |_) || |     | |_| |
 | |  | | |  __/   ___) ||  _ < | |___  |  _  |
 |_|  |_| |_|     |____/ |_| \_\ \____| |_| |_|  (TM)

*****************************************************************

        Release 2 1D John F  Collins, Biocomputing Research Unit
        Copyright (c) 1993, 1994, 1995  University of Edinburgh, U.K
                Distribution rights by IntelliGenetics, Inc

MPsrch_nn   n.a. - n.a. database search, using Smith-Waterman algorithm

Run on.           Thu Dec 18 18:14 27 1997;  MasPar time 401 48 Seconds
                                             981.565 Million cell updates/sec

Tabular output not generated

Title:            >US-08-826-361A-1
Description.      (1-1434) from US08826361A seq
Perfect Score:    1434
N.A. Sequence.        1 ATGAATTACAGCATTCCCAG.........AGAACCCACAGTCTCAGTGA 1434
     Comp:              TACTTAATGTCGTAAGGGTC.........TCTTGGGTGTCAGAGTCACT

Scoring table:    TABLE default
                  Gap 6

Nmatch   STD :    Dbase 0;  Query 0

Searched.         359085 seqs, 137405154 bases x 2

Post-processing   Minimum Match 0%
                  Listing first 45 summaries

Database:         EST-STS-THREE
                    1:EST199 2.EST200 3.EST201 4:EST202 5:EST203 6.EST204
                    7:EST205 8:EST206 9:EST207 10:EST208 11:EST209 12:EST210
                    13:EST211 14:EST212 15:EST213 16:EST214 17.EST215
                    18.EST216 19.EST217 20:EST218 21:EST219 22:EST220
                    23.EST221 24.EST222 25.EST223 26 EST224 27:EST225
                    28:EST226 29 EST227 30.EST228 31 EST229 32 EST230
                    33:EST231 34.EST232 35:EST233 36:STS1 37:STS2 38:STS3
                    39:STS4 40:STS5 41:STS6 42:STS7 43:STS8 44:STS9 45:STS10
                    46:STS11 47:STS12 48:STS13
Database:         EST-STS-FOUR
                    49:gnEST1 50.gnEST2 51.gnEST3 52:gnEST4 53:gnEST5
                    54.gnEST6 55:gnEST7 56:gnEST8 57:gnEST9 58:gnEST10
                    59:gnEST11 60:gnEST12 61.gnEST13 62:gnEST14 63:gnEST15
                    64:gnEST16 65:gnEST17 66:gnEST18 67:gnEST19 68:gnEST20
                    69:gnEST21 70.gnEST22 71.gnEST23 72:gnEST24 73:gnEST25
                    74:gnSTS 75.enEST1 76:enEST2 77:enEST3 78:enEST4
                    79:enEST5 80:enEST6 81:enEST7 82:enEST8 83:enEST9
                    84:enEST10 85:enEST11 86:enEST12 87:enEST13 88:enEST14
                    89:enEST15 90.enEST16 91:enEST17 92:enEST18 93:enEST19
                    94:enEST20 95:enEST21 96:enEST22 97:enEST23 98:enEST24
                    99:enEST25 100.enEST26 101:enEST27 102:enEST28
                    103:enEST29 104:enEST30 105:enEST31 106:enSTS 107:ueEST1
                    108:ueEST2

Statistics:       Mean 11.346;  Variance 2.040;  scale 5.562

       Pred. No. is the number of results predicted by chance to have a
       score greater than or equal to the score of the result being printed,
       and is derived by analysis of the total score distribution.

                                SUMMARIES

                      %
Result              Query
```

```
  No.  Score  Match Length DB  ID            Description            Pred. No.
-------------------------------------------------------------------------------
    1    53    3.7   2202 44  G28564        human STS SHGC-35295.  4.10e-55
c   2    52    3.6    390 96  MM7156        ma83b04.r1 Soares mou  3.29e-53
c   3    40    2.8    593 28  AA023625      mh80b06.r1 Soares mou  2 92e-31
    4    26    1.8    280 25  N59515        yy57g05 r1 Soares fet  5.81e-09
    5    24    1 7    156 46  DM36D7S       D  melanogaster STS d  2 91e-06
    6    24    1.7    536 97  MM94013       mb96d08.r1 Soares mou  2.91e-06
c   7    23    1.6    480 16  AA196298      zp95c10.s1 Stratagene  5.58e-05
    8    21    1.5    285 26  SSZ84056      S.scrofa mRNA; expres  1.45e-02
c   9    21    1.5    314 44  G27335        human STS CHLC ATA70B  1.45e-02
   10    21    1.5    320 90  HSW5605       ze02b04.r1 Soares fet  1.45e-02
   11    21    1.5    320  2  W95560        ze02b04 r1 Soares fet  1.45e-02
c  12    22    1.5    326 85  HS1164487     111371.seq.F Human fe  9.58e-04
c  13    22    1.5    326 69  AA249798      111371.seq.F Human fe  9 58e-04
   14    21    1.5    370 55  AA097671      mn85b08.r1 Stratagene  1 45e-02
   15    21    1.5    376 48  HUMUT552B     Human STS UT552, 3' p  1.45e-02
   16    22    1.5    382 30  AA057786      zf15e04.s1 Soares fet  9.58e-04
   17    21    1.5    384 83  HS1153545     zs38g06.r1 Soares NhH  1.45e-02
   18    21    1.5    384 66  AA235237      zs38g06.r1 Soares NhH  1.45e-02
c  19    21    1.5    387 15  AA132901      zp93g06 s1 Stratagene  1 45e-02
c  20    21    1.5    394  4  AA114039      zm29e11.r1 Stratagene  1.45e-02
c  21    22    1.5    399 21  AA021604      ze67c04.s1 Soares ret  9.58e-04
   22    21    1 5    422  7  AA126160      zm78b03.r1 Stratagene  1.45e-02
   23    22    1.5    424 89  HSW0986       ze08h09 r1 Soares fet  9.58e-04
   24    22    1.5    424  2  W96098        ze08h09.r1 Soares fet  9.58e-04
   25    21    1.5    426 65  AA233020      zr46a08 r1 Soares NhH  1.45e-02
c  26    21    1.5    443 72  AA253566      CrEST.158 pSKIIminusC  1 45e-02
c  27    21    1.5    445 81  HS1146141     nc27g02.r1 NCI CGAP P  1.45e-02
   28    22    1.5    452 28  AA024507      ze76d08.s1 Soares fet  9.58e-04
   29    21    1.5    466 77  AT60111       11379 Arabidopsis tha  1.45e-02
c  30    21    1.5    485 55  AA096657      mo03h09.r1 Stratagene  1.45e-02
c  31    22    1.5    498 11  AA179181      zp46b09.r1 Stratagene  9.58e-04
c  32    22    1.5    501 83  HS1151180     zr69e09 r1 Soares NhH  9.58e-04
c  33    22    1.5    501 65  AA233250      zr69e09.r1 Soares NhH  9.58e-04
c  34    21    1.5    538 103  MM1139113     mv68c05.r1 Soares mou  1.45e-02
c  35    21    1.5    538 49  AA220605      mv68c05.r1 Soares mou  1.45e-02
c  36    22    1.5    545 19  AA202284      LD02433 5prime LD Dro  9.58e-04
c  37    21    1.5    546 22  AA211579      zn56b11.s1 Stratagene  1 45e-02
c  38    21    1.5    565 20  AA208970      mw72g12 r1 Soares mou  1 45e-02
c  39    21    1.5    565 103  MMAA81339     mw72g12.r1 Soares mou  1.45e-02
c  40    21    1.5    578 19  AA205532      zq07a04 s1 Stratagene  1 45e-02
   41    21    1.5    586  7  AA123429      mp91a08 r1 Soares 2Nb  1.45e-02
   42    21    1.5    586 58  AA123429      mp91a08 r1 Soares 2Nb  1 45e-02
c  43    21    1.5    617 35  AA220111      mx36b03 r1 Barstead m  1.45e-02
c  44    21    1.5    667 52  AA144507      mt54g08.r1 Stratagene  1 45e-02
c  45    21    1.5    667 17  AA144507      mt54g08 r1 Stratagene  1 45e-02

                              ALIGNMENTS

RESULT      1
LOCUS       G28564        2202 bp    DNA              STS       01-JUL-1996
DEFINITION  human STS SHGC-35295.
ACCESSION   G28564
NID         g1408379
KEYWORDS    STS sequence, primer; sequence tagged site.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae, mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata, Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2202)
  AUTHORS   Myers,R.M.
  JOURNAL   Unpublished (1996)
COMMENT
            Contact  Richard M. Myers
            Stanford Human Genome Center (SHGC)
            Stanford University School of Medicine
            Department of Genetics, M-344, Stanford, CA 94305, USA
            Tel. 4157259687
            Fax: 4157259689
            Email: myers@shgc.stanford.edu
            Primer A: AATTGAACTCACTGCTTTTGAGG
            Primer B: CAGCTTCAGGTATCATCTTCACC
            STS size: 206
            PCR Profile:
                    Initial incubation: 94 degrees C for 90 seconds

                    Denaturation.         94 degrees C for 15 seconds
                    Annealing.            62 degrees C for 23 seconds
                    Polymerization:       72 degrees C for 30 seconds
                    PCR Cycles:           30
                    Thermal Cycler:       Perkin Elmer 9600
            Protocol:
                    Template:         25 ng
                    Primer:           each 1 uM
                    dNTPs:            each 200 uM
                    Taq Polymerase:   0.05 units/ul
                    Total Vol:        10 ul

            Buffer:
                    MgCl2:            2.5 mM
                    KCl:              50  mM
                    Tris-HCl:         20  mM
                    pH:               8.3

            Prepared with primer pairs provided by Sandoz, derived from M29960
            -- Washington University/Merck EST sequence
FEATURES             Location/Qualifiers
     source          1..2202
                     /organism "Homo sapiens"
                     /map="12"
     STS             1926..2131
     primer_bind     1926..1948
     primer_bind     complement(2109..2131)
BASE COUNT      720 a    442 c    439 g    601 t
ORIGIN

  Query Match              3.7%;  Score 53;  DB 44;  Length 2202;
  Best Local Similarity  61.9%;  Pred. No. 4.10e-55;
  Matches  138;  Conservative    0;  Mismatches   85;  Indels    0;  Gaps    0;

Db     368 aggaccaaataaaatttttgatctttacgtatgtatttgaaacaaagcatcaggacttca 427
           |||  |||| | || |    | |   | ||||   || ||   | ||    | ||||| |||   |||
Qy     261 AGGTTCAAAGAGGGATGCTCACTTTTGGGCTGTCTGCAGGGATTACGGATGGGATATCA 320

Db     428 ttatggagcagtaacttgtgaaggctgcaaaggatttttaaaagaagcatccgaaaaaa 487
             |||||||         | ||||||||| || || |   ||||||||||||||||| | |  | |
Qy     321 CTATGGAGTCTGGTCGTGTGAAGGATGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACA 380

Db     488 tttagtatattcatgtcgaggatcaaaggattgtattattaataagcaccaccgaaaccg 547
           |    |  |||    |||| ||    ||||  | ||||   ||   ||||   |||  || ||  |
Qy     381 TAATGATTATATTTGTCCAGCTACAAATCAGTGTACAATCGATAAAAACCGGCGCAAGAG 440

Db     548 ctgtcaatactgcaggttacagagatgtattgcgtttggaatg 590
           ||| ||    |||| |   | | ||   |||     |    | ||||||
Qy     441 CTGGCAGGGCTGCCGACTTCGGAAGTGTTACGAAGTGGGAATG 483


RESULT      2
ID    MM7156      standard; RNA; EST; 390 BP.
AC    W13715,
NI    g1287752
DT    30-APR-1996 (Rel. 47, Created)
DT    04-MAR-1997 (Rel. 51, Last updated, Version 2)
DE    ma83b04.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone 317263 5'
DE    similar to gb:L13740 EARLY RESPONSE PROTEIN NAK1 (HUMAN);
KW    EST.
OS    Mus musculus (house mouse)
OC    Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
OC    Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
OC    Mus.
RN    [1]
RP    1-390
RA    Marra M., Hillier L., Allen M., Bowles M., Dietrich N.,
RA    Dubuque T., Geisel S., Kucaba T., Lacy M., Le M., Martin J.,
RA    Morris M., Schellenberg K., Steptoe M., Tan F., Underwood K.,
RA    Moore B., Theising B., Wylie T., Lennon G., Soares B., Wilson R.,
RA    Waterston R.;
PT    "The WashU-HHMI Mouse EST Project";
RL    Unpublished.
CC    Contact: Marra M/Mouse EST Project WashU-HHMI Mouse EST Project
CC    Washington University School of MedicineF 4444 Forest Park Parkway,
CC    Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810
CC    Email: mouseest@watson.wustl.edu This clone is available
CC    royalty-free through LLNL ; contact the IMAGE Consortium
CC    (info@image.llnl.gov) for further information. Seq primer: ETPrimer
CC    High quality sequence stop: 385.
FH    Key             Location/Qualifiers
FH
FT    source          1..390
FT                    /organism="Mus musculus"
FT                    /note="Vector: pT7T3D (Pharmacia) with a modified
FT                    polylinker; Site_1. Not I, Site_2. Eco RI; 1st strand cDNA
FT                    was primed with a Not I - oligo(dT) primer [5'
FT                    TGTTACCAATCTGAAGTGGGAGCGGCCGCATTTTTTTTTTTTTTTTTTT 3'],
FT                    double-stranded cDNA was size selected, ligated to Eco RI
FT                    adapters (Pharmacia), digested with Not I and cloned into
FT                    the Not I and Eco RI sites of a modified pT7T3 vector
FT                    (Pharmacia). Library went through one round of
FT                    normalization to a Cot = 5. Library constructed by Bento
FT                    Soares and M.Fatima Bonaldo. RNA was kindly provided by Dr

FT                    Minoru Ko (Wayne State University)."
FT                    /clone="317263"
FT                    /clone_lib="Soares mouse p3NMF19.5"
FT                    /dev_stage="19.5 dpc total fetus"
FT                    /lab_host="DH10B (ampicillin resistant)"
FT    mRNA            <1..>390
SQ    Sequence 390 BP; 70 A; 106 C; 114 G; 100 T; 0 other;

  Query Match              3.6%;  Score 52;  DB 35;  Length 390;
  Best Local Similarity  62.7%;  Pred. No. 3.29e-53;
  Matches  129;  Conservative    0;  Mismatches   76;  Indels    0;  Gaps    0;

Db      80 cttaacaatcccgacactgagacacttctgaaatctgcagtactcacatcgatttcacg 139
           ||| ||||| || ||     |   ||||||| ||| || ||||   ||| || |   || || ||
Qy     489 CTTCACCATTCCCACTTCGTAACACTTCCGAAGTCGGCAGCCCTGCCAGCTCTTGCCCCC 430

Db     140 tctcttgtccactgggcagtttttatttgcccggcaaacatatttgcatttttctgcac 199
               | || || | ||   || |   ||   | ||   | |||| ||| |      || |    || |
Qy     429 GTTTTTATCGATTGTACACTGATTTTTAGCTGGACAAATATAATCATTATGTCCTTGAAT 370

Db     200 cgtttctcttgaagaagcccttgcaacccctcgcaggtgcctgactcgtaatgctgacaaac 259
               |||| || || ||| |||| || || || || |   | |||||| ||||| |   |   |
Qy     369 GCTTCTTTTAAAAAAGCCCTTACATCCTTCACACCACCAGACTCCATACTCATATCCCGA 310

Db     260 agcgttgtcgccgcagacggcaca 283
             ||||  |||| ||| || || ||
Qy     309 TGCGTAATCGCTGCAGACAGCCCA 286


RESULT      3
LOCUS       AA023625       593 bp    mRNA             EST       21 JAN 1997
DEFINITION  mh80b06.r1 Soares mouse placenta 4NbMPI3.5 14.5 Mus musculus cDNA
            clone 457235 5' similar to gb:M12674 ESTROGEN RECEPTOR (HUMAN);
            gb:M38651 Mouse estrogen receptor mRNA, complete cds (MOUSE);.
ACCESSION   AA023625
NID         g1487542
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 593)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:274123
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 492.
FEATURES             Location/Qualifiers
     source          1..593
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGGAAATTTTTTTTTTTTTTTTTTTTTT
                     T 3']; double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization, and was
                     constructed by Bento Soares and M.Fatima Bonaldo."
                     /clone="457235"
                     /clone_lib="Soares mouse placenta 4NbMP13.5 14.5"
                     /sex="unknown"
                     /dev_stage="adult"
                     /lab_host="DH10B"
     mRNA            <1..>593
BASE COUNT      144 a    165 c    119 g    165 t
ORIGIN

  Query Match             2.8%;  Score 40;  DB 28;  Length 593;
  Best Local Similarity  74.4%,  Pred. No. 2.92e-31;
  Matches   61;  Conservative    0;  Mismatches  21;  Indels    0;  Gaps    0;

Db     174 tctttgcccagttgatcatatgaaccagctccctatctgctaggttggtcaataagccca 233
           |||| |||||| ||||||| || |||| |||| | || || |  |||||||    |   ||
Qy     784 TCTTGGCCCAGCTGATCATGTGTACCAACTCCTTGTCGGCCAACTTGGTCAGGGACATCA 725

Db     234 tcattgaggcttcactgaaggg 255
           |||| ||||| ||  |||||||
Qy     724 TCATGGAGGCCTCGGTGAAGGG 703


RESULT      4
LOCUS       N59515          280 bp    mRNA            EST        28-JAN-1997
DEFINITION  yv57g05.r1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone
            246872 5' similar to SW LRH1_MOUSE P45448 LIVER RECEPTOR HOMOLOG :.
ACCESSION   N59515
NID         g1203405
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 280)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E.,
            Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Insert Length: 729   Std Error: 0.00
            Seq primer: T7
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1..280
                     /organism="Homo sapiens"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Pac I - oligo(dT) primer
                     [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
                     /clone="246872"
                     /clone_lib="Soares fetal liver spleen 1NFLS"
                     /sex="male"
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
     mRNA            <1..>280
BASE COUNT       80 a     45 c     75 g     61 t     19 others
ORIGIN

  Query Match             1.8%;  Score 26;  DB 25;  Length 280;
  Best Local Similarity  54.0%,  Pred. No. 5.81e-09;
  Matches   94;  Conservative    0;  Mismatches  80;  Indels    0;  Gaps    0;

Db      26 tcngggtancattangggctcctcanctgtgnaagctgnaaggggatttttaagcgnaca 85
           || || || || || ||  ||         |||| | | || ||||  ||||||||  | |
Qy     310 TCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATGTAAGGCCTTTTTAAAACAAGC 369

Db      86 gtccaaaataataaaaggtacacntgnatataaaaccagnactgccgaattgncaaaacn 145
            | |||     ||||    || |  ||     |    |    |   | ||   ||| |  ||||
Qy     370 ATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCAGTGTACAATCGATAAAAAC 429

Db     146 cagagaaagcgttgtccttactgtcgttttcaaaaatgtctaagtgttggaatg 199
           | | | ||| | || |     ||| ||  |||  || |||      || ||||||
Qy     430 CGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTACGAAGTGGGAATG 483


RESULT      5
LOCUS       DM36D7S         156 bp    DNA             STS        16-APR-1996
DEFINITION  D. melanogaster STS determined from European Mapping Project
            cosmid.
ACCESSION   Z70880
NID         g1263754
KEYWORDS    sequence tagged site.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Arthropoda;
            Tracheata; Insecta; Pterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 156)
  AUTHORS   European Drosophila Mapping,Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-APR-1996) Michael Ashburner, Department of Genetics,
            Downing St., Cambridge CB2 3EH, England
COMMENT     STS_name = Dm36D7S
            clone_name = 36D7
            STS_from_promoter = SP6
            vector_class = cosmid, Lorist 6
            origin_of_clone = Oregon-R
            in_situ_site_primary - 98C
            BLAST_program - BLASTN
            database_searched = EMBL
            database_version - 45.0 and updates till date_of_search
            date_of_search - 08-01-1996
            BLAST_program - BLASTX
            database_searched = SWISSPROT
            database_version   32.0
            date_of_search - 15-12-1995
FEATURES             Location/Qualifiers
     source          1..156
                     /organism="Drosophila melanogaster"
                     /strain="Oregon-R"
                     /clone="36D7"
BASE COUNT       28 a     37 c     30 g     27 t     34 others
ORIGIN

  Query Match             1.7%,  Score 24,  DB 36,  Length 156,
  Best Local Similarity  19.6%;  Pred. No. 2.91e-06;
  Matches     9;  Conservative   26;  Mismatches  11;  Indels   0;  Gaps   0;

Db      21 [illegible] 66
           |: ::::| :||:::: |:::::::: : | :: : :: ||| ::
Qy     597 GCCCCGAGTGCGGGAGCTGCTGCTGGACGCCCTGAGCCCCGAGCAG 642


RESULT    6
ID   MM94013    standard; RNA; EST; 536 BP.
AC   W20940;
NI   g1297857
DT   08-MAY-1996 (Rel. 47, Created)
DT   05-MAR-1997 (Rel. 51, Last updated, Version 2)
DE   mb96d08.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone 337263 5'
DE   similar to gb:M64497 APOLIPOPROTEIN AI REGULATORY PROTEIN-1
DE   (HUMAN); gb:U07635 Mus musculus ovalbumin upstream promoter
DE   transcription factor (MOUSE);.
KW   EST.
OS   Mus musculus (house mouse)
OC   Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
OC   Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
OC   Mus.
RN   [1]
RP   1-536
RA   Marra M., Hillier L., Allen M., Bowles M., Dietrich N.,
RA   Dubuque T., Geisel S., Kucaba T., Lacy M., Le M., Martin J.,
RA   Morris M., Schellenberg K., Steptoe M., Tan F., Underwood K.,
RA   Moore B., Theising B., Wylie T., Lennon G., Soares B., Wilson R.,
RA   Waterston R.;
RT   "The WashU-HHMI Mouse EST Project";
RL   Unpublished.
CC   Contact: Marra M/Mouse EST Project WashU-HHMI Mouse EST Project
CC   Washington University School of MedicineP 4444 Forest Park Parkway,
CC   Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810
CC   Email: mouseest@watson.wustl.edu This clone is available
CC   royalty-free through LLNL ; contact the IMAGE Consortium
CC   (info@image.llnl.gov) for further information. MGI:218663 Seq
CC   primer: ETPrimer High quality sequence stop: 385.
FH   Key             Location/Qualifiers
FH
FT   source          1..536
FT                   /organism="Mus musculus"
FT                   /note="Vector: pT7T3D (Pharmacia) with a modified
FT                   polylinker; Site_1: Not I; Site_2: Eco RI. 1st strand cDNA
FT                   was primed with a Not I - oligo(dT) primer [5'
FT                   TGTTACCAATCTGAAGTGGGAGCGGCCGCATTTTTTTTTTTTTTTTTTT 3'],
FT                   double-stranded cDNA was size selected, ligated to Eco RI
FT                   adapters (Pharmacia), digested with Not I and cloned into
FT                   the Not I and Eco RI sites of a modified pT7T3 vector
FT                   (Pharmacia). Library went through one round of
FT                   normalization to a Cot = 5. Library constructed by Bento
FT                   Soares and M.Fatima Bonaldo. RNA was kindly provided by Dr

FT                   Minoru Ko (Wayne State University)."
```

```
FT                   /clone="337263"
FT                   /clone_lib="Soares mouse p3NMF19.5"
FT                   /dev_stage="19.5 dpc total fetus"
FT                   /lab_host="DH10B (ampicillin resistant)"
FT   mRNA            <1..>536
SQ   Sequence 536 BP; 113 A; 175 C; 164 G; 84 T; 0 other;

  Query Match             1.7%,  Score 24,  DB 37,  Length 536;
  Best Local Similarity  66.7%;  Pred. No. 2.91e-06;
  Matches    42;  Conservative    0;  Mismatches   24;  Indels    0;  Gaps    0;

Db     205 [illegible] 264
           ||| | |||||  |  |||  |||||    ||||||  |||||| ||     ||||    |
Qy     412 TGTACAATCCATAAAAACCGGGGCAAGAGCTGCCAGCCCTGCCCACTTCGGAAGTGTTAC 471

Db     265 aaagtgggcatg 276
            ||||||| |||
Qy     472 GAAGTGGGAATG 483


RESULT    7
LOCUS       AA196298      483 bp    mRNA            EST       22-JAN-1997
DEFINITION  zp95c10.s1 Stratagene muscle 937209 Homo sapiens cDNA clone 627954
            3'.
ACCESSION   AA196298
NID         g1791889
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 480)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E.,
            Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40M13 fwd. from Amersham
            High quality sequence stop: 204.
FEATURES             Location/Qualifiers
     source          1..480
                     /organism="Homo sapiens"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI; Cloned unidirectionally. Primer: Oligo dT. Skeletal
                     muscle from patient with malignant hyperthermia.  Average
                     insert size: 1.0 kb, Uni-ZAP XR Vector: ~5' adaptor
                     sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor sequence: 5'
                     CTCGAGTTTTTTTTTTTTTTTTTTTT 3'"
                     /clone="627954"
                     /clone_lib="Stratagene muscle 937209"
                     /dev_stage="adult"
                     /lab_host="SOLR (kanamycin resistant)"
     mRNA            complement(<1..>480)
BASE COUNT       84 a    151 c    133 g    107 t      5 others
ORIGIN

  Query Match             1.6%;  Score 23;  DB 16;  Length 480;
  Best Local Similarity  69.0%;  Pred. No. 5.58e-05;
  Matches    40;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Db     203 [illegible] 260
           |||||| ||| || |||  || | | ||  ||| | ||||  ||||| || || | ||
Cp     250 CAGGGCTGGCGTAACGGTTCCCACTAACCTTCCTTTTCAGTGTCTCTCTGTTTACAGG 193


RESULT     8
LOCUS       SSZ84056       285 bp    RNA              EST       21-JAN-1997
DEFINITION  S.scrofa mRNA: expressed sequence tag (5', clone c13c02).
ACCESSION   Z84056
NID         g1806383
KEYWORDS    EST; expressed sequence tag, UDP glucose 6-dehydrogenase.
SOURCE      pig.
  ORGANISM  Sus scrofa
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Artiodactyla; Suiformes; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 285)
  AUTHORS   Winteroe,A.K. and Fredholm,M.
  TITLE     Evaluation and characterization of a porcine small intestine cDNA
            library
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 285)
  AUTHORS   Winteroe,A.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JAN-1997) Winteroe A.K., The Royal Veterinary and
            Agricultural University, Department of Animal Science and Animal
            Health, Division of Animal Genetics, Bulowsvej 13, 1870
            Frederiksberg C, DENMARK
FEATURES             Location/Qualifiers
     source          1..285
                     /organism="Sus scrofa"
                     /tissue_type="small intestine"
                     /clone_lib="directionally cloned cDNA in XL1-blue MPF'"
                     /clone="c13c02"
     CDS             151..>285
                     /gene="UDP glucose 6-dehydrogenase"
                     /note="by similarity to bovine sequence"
                     /codon_start=1
                     /db_xref="PID:e293098"
                     /db_xref="PID:g1806384"
                     /translation="MFEIKKICCIGAGYVGGPTCSVIALKCPEIRVTVVDVNESRINA
                     W"
BASE COUNT       66 a     57 c     99 g     62 t      1 others
ORIGIN

  Query Match             1.5%;  Score 21;  DB 26;  Length 285;
  Best Local Similarity  75.0%;  Pred. No. 1.45e 02;
  Matches    30;  Conservative    0;  Mismatches  10;  Indels    0;  Gaps    0;

Db      52 [illegible] 91
            |||||| | |||||| ||| |  |||| ||||    ||||
Qy     601 CGAGTGCGGGAGCTGCTGCTGGACGCCCTGAGCCCCGACC 640


RESULT     9
LOCUS       G27335         314 bp    DNA              STS       19-JUN-1996
DEFINITION  human STS CHLC.ATA70B03.P68256 clone ATA70B03.
ACCESSION   G27335
NID         g1381374
KEYWORDS    STS sequence; primer; sequence tagged site
SOURCE      human vector=pJCP1 host=E.coli dut-ung- (DH10B) Marker Selected
            genomic DNA prepared from XY individual of French nationality.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 314)
  AUTHORS   Murray,J., Sheffield,V., Weber,J.L., Duyk,G. and Buetow,K.H.
  TITLE     Cooperative Human Linkage Center
  JOURNAL   Unpublished (1995)
COMMENT     Synonyms: ATA70B03, CHLC.ATA70B03.P68207
            Contact: Dr. Jeffrey C. Murray
            UofI
            The University of Iowa
            Department of Pediatrics, Iowa City, IA 52242, USA
            Tel: (319) 356-3508
            Fax: (319) 356-3347
```

```
            Email: jeff-murray@uiowa.edu

            Primer A: GACTTTGTAAGGTTTCTTTGTGG
            Primer B: GGCTGAGGTGGGAAGATC
            STS size: 173
            PCR Profile:
                     denature:    30 seconds at 94 degrees C
                     annealing:   75 seconds at 55 degrees C
                     extension:   15 seconds at 72 degrees C
                     PCR cycles:  27
                     extension:    6 minutes at 72 degress C
            Protocol:
                     Template:          30ng genomic DNA
                     Primer:            each 1.5 pmole
                     dNTPs:             each 200 uM
                     Taq Polymerase:    0.3 units
                     Total Vol:         10 ul

            Buffer:
                     MgCl2:  1.5mM
                     KCl:     50mM
                     Tris.    10mM
                     pH:     8.3.
FEATURES             Location/Qualifiers
     source          1..314
                     /organism="Homo sapiens"
     STS             56..228
     primer_bind     56..78
     primer_bind     complement(211..228)
BASE COUNT       70 a     57 c     70 g    102 t     15 others
ORIGIN

  Query Match             1.5%,  Score 21,  DB 44,  Length 314;
  Best Local Similarity  70.6%,  Pred. No. 1.45e-02;
  Matches    36,  Conservative    0,  Mismatches  15;  Indels    0;  Gaps    0;

Db     130 [illegible] 180
            ||| ||    ||||| | |||||||  || | ||| | || | || |||||
Cp     680 ACATGGGGCGGCTCAGCCTCCAGGAGGGTGAGCACTAGCTGCTGGGGCTC 630


RESULT   10
ID   HSW5605    standard; RNA; EST; 320 BP.
AC   W95560;
NI   g1425468
DT   18-JUL-1996 (Rel. 48, Created)
DT   28-FEB-1997 (Rel. 51, Last updated, Version 2)
DE   ze02b04.r1 Soares fetal heart NbHH19W Homo sapiens cDNA clone
DE   357775 5' similar to PIR:JC4014 JC4014 steroid hormone-nuclear
DE   receptor NER - human ;.
KW   EST.
OS   Homo sapiens (human)
OC   Eukaryotae; mitochondrial eukaryotes, Metazoa; Chordata;
OC   Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   1-320
RA   Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,
RA   Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,
RA   Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,
RA   Trevaskis E., Waterston R., Williamson A., Wohldmann P., Wilson R.;
RT   "The WashU-Merck EST Project";
RL   Unpublished.
CC   Contact: Wilson RK WashU-Merck EST Project Washington University
CC   School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis,
CC   MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:
CC   est@watson.wustl.edu This clone is available royalty-free through
CC   LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for
CC   further information. Insert Length: 810 Std Error: 0.00 Seq primer:
CC   mob.REGA+ET High quality sequence stop: 293.
FH   Key             Location/Qualifiers
FH
FT   source          1..320
FT                   /organism="Homo sapiens"
FT                   /note="Organ: heart, Vector: pT7T3D (Pharmacia) with a
FT                   modified polylinker. Site_1  Not I; Site_2. Eco RI, 1st
FT                   strand cDNA was primed with a Not I - oligo(dT) primer [5'
FT                   TGTTACCAATCTGAAGTGGGAGCGGCCGCATCTTTTTTTTTTTTTTTTTT 3'],
FT                   double-stranded cDNAwas size selected, ligated to Eco RI
FT                   adapters (Pharmacia), digested with Not I and cloned into
FT                   the Not I and Eco RI sites of a modified pT7T3 vector
FT                   (Pharmacia). Library went through one round of
FT                   normalization to a Cot = 5. Library constructed by M.Fatim
 a
FT                   Bonaldo. This library was constructed from the same fetus
FT                   as the fetal lung library, Soares fetal lung NbHL19W."
FT                   /clone="357775"
FT                   /clone_lib="Soares fetal heart NbHH19W"
FT                   /sex="unknown"
FT                   /dev_stage="19 weeks"
FT                   /lab_host="DH10B (ampicillin resistant)"
FT   mRNA            <1..>320
SQ   Sequence 320 BP; 76 A; 83 C; 83 G; 76 T; 2 other;

  Query Match             1.5%;  Score 21;  DB 90;  Length 320;
  Best Local Similarity  66.7%;  Pred. No. 1.45e-02;
  Matches   52;  Conservative    0;  Mismatches  25;  Indels    1;  Gaps    1;

Db      44 gctaaacagctacccggcttcctgcagctcagccggga-gaccagattgccctgctgaag 102
           || ||  || | ||||||||  || |||||||||| |   |||||  |     ||  || ||
Qy     778 GCCAAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAAGTGCGGCTCTTGGAG 837

Db     103 acctctgcgatcgaggtg 120
           | || |  ||| ||||||
Qy     838 AGCTGTTGGATGGAGGTG 855


RESULT    11
LOCUS       W95560        320 bp    mRNA            EST       25-NOV-1996
DEFINITION  ze02b04.r1 Soares fetal heart NbHH19W Homo sapiens cDNA clone
            357775 5' similar to PIR:JC4014 JC4014 steroid hormone nuclear
            receptor NER - human ;.
ACCESSION   W95560
NID         g1425468
KEYWORDS    EST
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo
REFERENCE   1  (bases 1 to 320)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 810   Std Error: 0.00
            Seq primer: mob.REGA+ET
            High quality sequence stop: 293.
FEATURES             Location/Qualifiers
     source          1..320
                     /organism="Homo sapiens"
                     /note="Organ: heart; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I, Site_2: Eco RI, 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCATCTTTTTTTTTTTTTTTTTT 3'],
```

```
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library constructed by
                     M.Fatima Bonaldo. This library was constructed from the
                     same fetus as the fetal lung library, Soares fetal lung
                     NbHL19W."
                     /clone="357775"
                     /clone_lib="Soares fetal heart NbHH19W"
                     /sex="unknown"
                     /dev_stage="19 weeks"
                     /lab_host="DH10B (ampicillin resistant)"
     mRNA            <1..>320
BASE COUNT       76 a     83 c     83 g     76 t      2 others
ORIGIN

  Query Match             1.5%;  Score 21;  DB 2;  Length 320;
  Best Local Similarity  66.7%;  Pred. No. 1.45e-02;
  Matches   52;  Conservative    0;  Mismatches  25;  Indels    1;  Gaps    1;

Db      44 gctaaacagctacccggcttcctgcagctcagccggga-gaccagattgccctgctgaag 102
           || ||  || | ||||||||  || |||||||||| |   |||||  |     ||  || ||
Qy     778 GCCAAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAAGTGCGGCTCTTGGAG 837

Db     103 acctctgcgatcgaggtg 120
           | || |  ||| ||||||
Qy     838 AGCTGTTGGATGGAGGTG 855


RESULT    12
ID   HS1164487  standard; RNA; EST; 326 BP.
AC   AA249798;
NI   g1881068
DT   13-MAR-1997 (Rel. 51, Created)
DT   13-MAR-1997 (Rel. 51, Last updated, Version 1)
DE   111371.seq.F Human fetal heart, Lambda ZAP Express Homo sapiens
DE   cDNA 5'.
KW   EST.
OS   Homo sapiens (human)
OC   Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
OC   Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   1-326
RA   Liew C.C.;
RT   "cDNAs from human fetal heart (1997)";
RL   Unpublished.
CC   Contact: Liew CC Molecular Cardiology University of Toronto Banting
CC   Institute, 100 College St., Toronto, Ontario, M5G1L5 Tel:
CC   4169788758 Fax: 4169785650 Email: liewcc@utcc.utoronto.ca PCR
CC   PRimers FORWARD: 5' GCCAAGCTCGAAATTAACCCTCACTAAAGGG 3' BACKWARD: 5'
CC   CCAGTGAATTGTAATACGACTCACTATAGGGCG 3' Seq primer: 5'
CC   GAAATTAACCCTCACTAAAGGG 3'.
FH   Key             Location/Qualifiers
FH
FT   source          1..326
FT                   /organism="Homo sapiens"
FT                   /note="Vector: Lambda ZAP Express, Site_1: EcoRI; Site_2:
FT                   XhoI, mRNA was purified from human fetal hearts (8-10
FT                   weeks). cDNA was synthesized using a XhoI-Oligo dT
FT                   adaptor primer. EcoRI adaptors were ligated, followed by
FT                   digestion with XhoI, fordirectional cloning into
FT                   predigested lambda ZAP Express."
FT                   /clone_lib="Human fetal heart, Lambda ZAP Express"
FT                   /lab_host="E. coli XL1-Blue"
FT   mRNA            <1..>326
SQ   Sequence 326 BP; [illegible] A; 87 [illegible] other;

  Query Match             1.5%;  Score 22;  DB 85;  Length 326;
  Best Local Similarity  78.9%;  Pred. No. 9.58e-04;
  Matches   30;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Db     257 cgtctgcagctcattatctgggacaaggcggtaccca 294
```

```
                |:|| |:| ||  : |||  |||||:|:||||:||||
Cp     550 CGTCGGCACTTCTCTGTCTCCGCACAAGGCGGTACCCA 513


RESULT   13
LOCUS        AA249798       326 bp    mRNA                EST       11-MAR-1997
DEFINITION   111371.seq.F Human fetal heart, Lambda ZAP Express Homo sapiens
             cDNA 5'.
ACCESSION    AA249798
NID          g1881068
KEYWORDS     EST.
SOURCE       human.
  ORGANISM   Homo sapiens
             Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
             Vertebrata; Eutheria; Primates; Catarrhini; Hominidae, Homo.
REFERENCE    1  (bases 1 to 326)
  AUTHORS    Liew,C.C.
  TITLE      cDNAs from human fetal heart (1997)
  JOURNAL    Unpublished (1997)
COMMENT
             Contact: Liew CC
             Molecular Cardiology
             University of Toronto
             Banting Institute, 100 College St., Toronto, Ontario, M5G1L5
             Tel: 4169788758
             Fax: 4169785650
             Email: liewcc@utcc utoronto.ca
             PCR PRimers
             FORWARD: 5' GCCAAGCTCGAAATTAACCCTCACTAAAGGG 3'
             BACKWARD: 5' CCAGTGAATTGTAATACGACTCACTATAGGGCG 3'
             Seq primer: 5' GAAATTAACCCTCACTAAAGGG 3'.
FEATURES             Location/Qualifiers
     source          1..326
                     /organism="Homo sapiens"
                     /note="Vector: Lambda ZAP Express; Site_1: EcoRI; Site_2:
                     XhoI; mRNA was purified from human fetal hearts (8-10
                     weeks).  cDNA was synthesized using a XhoI-Oligo dT
                     adaptor-primer.  EcoRI adaptors were ligated, followed by
                     digestion with XhoI, for directional cloning into
                     predigested lambda ZAP Express."
                     /clone_lib="Human fetal heart, Lambda ZAP Express"
                     /lab_host="E. coli XL1-Blue"
     mRNA            <1..>326
BASE COUNT       63 a     87 c     86 g     90 t
ORIGIN

  Query Match            1.5%;  Score 22;  DB 69;  Length 326;
  Best Local Similarity  78.9%;  Pred. No. 9.58e-04;
  Matches    30;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Db     257 cgtctgcagctcattatctgggcacaaggcggtaccca 294
           |:|| |||  :|  : |||  ||||:||||||||||||
Cp     550 CGTCGGCACTTCTCTGTCTCCGCACAAGGCGGTACCCA 513


RESULT   14
LOCUS        AA097671       370 bp    mRNA                EST       15-FEB-1997
DEFINITION   mn85b08.r1 Stratagene mouse Tcell 937311 Mus musculus cDNA clone
             550839 5'.
ACCESSION    AA097671
NID          g1643459
KEYWORDS     EST.
SOURCE       house mouse.
  ORGANISM   Mus musculus
             Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
             Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
             Mus.
REFERENCE    1  (bases 1 to 370)
  AUTHORS    Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
             Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
             Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
             Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
             Waterston,R.
  TITLE      The WashU-HHMI Mouse EST Project
  JOURNAL    Unpublished (1996)
COMMENT
             Contact: Marra M/Mouse EST Project
             WashU-HHMI Mouse EST Project
             Washington University School of MedicineP
             4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
             Tel: 314 286 1800
             Fax: 314 286 1810
             Email: mouseest@watson.wustl.edu
             This clone is available royalty-free through LLNL ; contact the
             IMAGE Consortium (info@image.llnl.gov) for further information.
             MGI:331631
             Seq primer: -28m13 rev1 ET from Amersham
             High quality sequence stop: 349.
FEATURES             Location/Qualifiers
     source          1..370
                     /organism="Mus musculus"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI; Cloned unidirectionally.  Primer: Oligo dT. M30 CD4+
                     cells. Average insert size: 1.0 kb; Uni ZAP XR Vector; -5'
                     adaptor sequence: 5' GAATTCGGCACGAG 3' -3' adaptor
                     sequence: 5' CTCGAGTTTTTTTTTTTTTTTTTT 3'"
                     /clone="550839"
                     /clone_lib="Stratagene mouse Tcell 937311"
                     /dev_stage="M30 CD4+ cells"
                     /lab_host="SOLR (kanamycin resistant)"
     mRNA            <1..>370
BASE COUNT       84 a    112 c    116 g     58 t
ORIGIN

  Query Match            1.5%;  Score 21;  DB 55;  Length 370;
  Best Local Similarity  75.6%;  Pred. No. 1.45e-02;
  Matches    31;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Db      19 agggaagtcggaatggaagtgctggcctgagactggccca 58
           |||  ||| |||||||   ||||||| |||  ||||| |||
Qy     221 AGGTTAGTGGGAACTGTTGTTTCAGTCCTGTTACTGGTCCA 261


RESULT   15
LOCUS        HUMUT552B      376 bp    DNA                 STS       25-MAR-1994
DEFINITION   Human STS UT552, 3' primer bind.
ACCESSION    L30016
NID          g465246
KEYWORDS     PCR primer; STS sequence; microsatellite DNA; microsatellite
             marker; sequence tagged site; tetranucleotide repeat.
SOURCE       Homo sapiens DNA.
  ORGANISM   Homo sapiens
             Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
             Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 376)
  AUTHORS    Gerken,S.C., Matsunami,N., Plaetke,R., Albertsen,H., Ballard,L.,
             Melis,R., Lawrence,E., Moore,M., Holik,P.R., Carlson,M., Zhao,X.,
             Robertson,M., Bradley,P., Elsner,T., Tingey,A., Lalouel,J.-M. and
             White,R.
  TITLE      Genetic and physical mapping of simple sequence repeat containing
             sequence tagged sites from the human genome
  JOURNAL    Unpublished (1994)
COMMENT      Submitted by: Utah Center for Human Genome Research University of
             Utah, Dept. of Human Genetics
             2160 Eccles Institute of Human Genetics
             Salt Lake City, UT  84112
             e-mail: sts@corona.med.utah.edu
             Primer A:  GGTCAAGGCTGCAGTGCAG
             Primer B:  AGCTGGGACCAGAGGCACA
             End to Label: Primer B
             PCR Profile:
             Initial Denaturation: 94C 300sec
             Cycles     Denaturation     Annealing        Extension 5          94
             C  10 sec.    62 C  10 sec.    72 C  20 sec. 30             94 C  10 sec.
               58 C  10 sec     72 C  20 sec. Mg++: 1.00 mM
             Gel: Acrylamide 7%, Formamide 32%, Urea 34%
             Alleles: 1.
FEATURES             Location/Qualifiers
     source          1..376
                     /organism="Homo sapiens"
                     /map="12"
     primer_bind     complement(86..104)
                     /evidence=experimental
BASE COUNT      123 a     75 c     89 g     84 t       5 others
ORIGIN

  Query Match             1.5%;  Score 21;  DB 48;  Length 376;
  Best Local Similarity  86.2%;  Pred. No. 1.45e-02;
  Matches    25;  Conservative    0;  Mismatches   4;  Indels   0;  Gaps   0;

Db      90 cctctggtcccagctactcaggaggctga 118
           ||||||||| |||| || ||||| |||||
Qy    1075 CCTCTGGTCACAGCGACCCAGGATGCTGA 1103


Search completed: Thu Dec 18 18:21:19 1997
Job time : 412 secs.
```

```
**********************************************************************
                         MPSRCH                                  (TM)
**********************************************************************

        Release 2 1D John F. Collins, Biocomputing Research Unit.
        Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
                 Distribution rights by IntelliGenetics, Inc

MPsrch_nn    n a  -  n a  database search, using Smith-Waterman algorithm

Run on:          Thu Dec 18 18:00:55 1997; MasPar time 774.54 Seconds
                                           950.371 Million cell updates/sec
Tabular output not generated

Title:           >US-08-826-361A-1
Description:     (1-1434) from US08826361A.seq
Perfect Score:   1434
N.A. Sequence:       1 ATGAATTACAGCATTCCCAG           AGAACCCACAGTCTCAGTGA 1434
    Comp.              TACTTAATGTCGTAAGGGTC           TCTTGGGTGTCAGAGTCACT

Scoring table:   TABLE default
                 Gap 6

Nmatch   STD :   Phase 0:  Query 0

Searched:        737717 seqs  256659330 bases x 2

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database:        EST-STS
                   1:EST1 2:EST2 3:EST3 4:EST4 5:EST5 6:EST6 7:EST7 8:EST8
                   9:EST9 10:EST10 11:EST11 12:EST12 13:EST13 14:EST14
                   15:EST15 16:EST16 17:EST17 18:EST18 19:EST19 20:EST20
                   21:EST21 22:EST22 23:EST23 24:EST24 25:EST25 26:EST26
                   27:EST27 28:EST28 29:EST29 30:EST30 31:EST31 32:EST32
                   33:EST33 34:EST34 35:EST35 36:EST36 37:EST37 38:EST38
                   39:EST39 40:EST40 41:EST41 42:EST42 43:EST43 44:EST44
                   45:EST45 46:EST46 47:EST47 48:EST48 49:EST49 50:EST50
                   51:EST51 52:EST52 53:EST53 54:EST54 55:EST55 56:EST56
                   57:EST57 58:EST58 59:EST59 60:EST60 61:EST61 62:EST62
                   63:EST63 64:EST64 65:EST65 66:EST66 67:EST67 68:EST68
                   69:EST69 70:EST70 71:EST71 72:EST72 73:EST73 74:EST74
                   75:EST75 76:EST76 77:EST77 78:EST78 79:EST79 80:EST80
                   81:EST81 82:EST82 83:EST83 84:EST84 85:EST85 86:EST86
                   87:EST87 88:EST88 89:EST89 90:EST90 91:EST91 92:EST92
                   93:EST93 94:EST94 95:EST95 96:EST96 97:EST97 98:EST98
                   99:EST99
Database:        EST-STS-TWO
                   100:EST100 101:EST101 102:EST102 103:EST103 104:EST104
                   105:EST105 106:EST106 107:EST107 108:EST108 109:EST109
                   110:EST110 111:EST111 112:EST112 113:EST113 114:EST114
                   115:EST115 116:EST116 117:EST117 118:EST118 119:EST119
                   120:EST120 121:EST121 122:EST122 123:EST123 124:EST124
                   125:EST125 126:EST126 127:EST127 128:EST128 129:EST129
                   130:EST130 131:EST131 132:EST132 133:EST133 134:EST134
                   135:EST135 136:EST136 137:EST137 138:EST138 139:EST139
                   140:EST140 141:EST141 142:EST142 143:EST143 144:EST144
                   145:EST145 146:EST146 147:EST147 148:EST148 149:EST149
                   150:EST150 151:EST151 152:EST152 153:EST153 154:EST154
                   155:EST155 156:EST156 157:EST157 158:EST158 159:EST159
                   160:EST160 161:EST161 162:EST162 163:EST163 164:EST164
                   165:EST165 166:EST166 167:EST167 168:EST168 169:EST169
                   170:EST170 171:EST171 172:EST172 173:EST173 174:EST174
                   175:EST175 176:EST176 177:EST177 178:EST178 179:EST179
                   180:EST180 181:EST181 182:EST182 183:EST183 184:EST184
                   185:EST185 186:EST186 187:EST187 188:EST188 189:EST189
                   190:EST190 191:EST191 192:EST192 193:EST193 194:EST194
                   195:EST195 196:EST196 197:EST197 198:EST198

Statistics.        Mean 11.322;  Variance 1.989;  scale 5.693

        Pred. No. is the number of results predicted by chance to have a
        score greater than or equal to the score of the result being printed,
        and is derived by analysis of the total score distribution.

                                  SUMMARIES

                     %
Result               Query
    No.   Score   Match Length DB  ID             Description            Pred. No.
-------------------------------------------------------------------------------
c     1     53     3.7    454  71  H82542       ys69h11.r1 Homo sapie  2.73e-56
c     2     52     3.6    390 149   W13715       ma83b04.r1 Soares mou  2.43e-54
c     3     52     3.6    403  74  H91890       ys81d07.r1 Homo sapie  2.43e-54
      4     51     3.6    412 191   AA157597     zo44a02.r1 Stratagene  2.12e-52
      5     48     3.3    509 120   W53545       md56a06.r1 Soares mou  1.20e-46
      6     48     3.3    536 147   AA003987     mg80b08.r1 Soares mou  1.20e-46
      7     41     2.9    552 120   W53477       md45d07.r1 Soares mou  1.17e-33
      8     38     2.6    365 181   AA092184     114009.seq.F Fetal he  2.53e-28
      9     38     2.6    392  26  R52304       yg79c04.r1 Homo sapie  2.53e-28
     10     38     2.6    504 195   AA049997     mj18e07.r1 Soares mou  2.53e-28
     11     36     2.5    566 120   W53511       mj45f07.r1 Soares mou  7.37e-25
     12     32     2.2    378 141   C08492       C.elegans cDNA clone   2.44e-18
     13     29     2.0    163   8  T83788       yd65a11.r1 Homo sapie  1.84e-13
     14     28     2.0    360  89  CELK105GXF   C.elegans cDNA clone   6.03e-12
     15     28     2.0    431 137   AA044343     zk54a08.r1 Soares pre  6.03e-12
     16     28     2.0    578 125   W74889       me32d11.r1 Soares mou  6.03e-12
     17     27     1.9    368 139   N44459       SW3C0A1324SK Brugia m  1.82e-10
     18     27     1.9    453   6  T75407       yd11c03.r1 Homo sapie  1.82e-10
     19     25     1.7    120 141   C09766       C.elegans cDNA clone   1.29e-07
     20     24     1.7    141  65  H28316       yl60e10.r1 Homo sapie  2.96e-06
     21     24     1.7    336 154   AA011708     ze10d04.r1 Soares fet  2.96e-06
c    22     24     1.7    340 154   AA011660     ze10d04.s1 Soares fet  2.96e-06
     23     25     1.7    360  89  CELK092E1F   C.elegans cDNA clone   1.29e-07
     24     25     1.7    382 118   W39474       zc76d10.r1 Pancreatic  1.29e-07
     25     25     1.7    365 146   T26966       NIB1014D11R Infant br  1.29e-07
     26     24     1.7    416  43  M78627       EST01014 Homo sapiens  2.96e-06
     27     24     1.7    466 154   AA019673     ze16f11.r1 Soares fet  2.96e-06
     28     24     1.7    487  63  H21142       ym51d05.r1 Homo sapie  2.96e-06
     29     24     1.7    535 121   W21340       mb52d08.r1 Soares mou  2.96e-06
     30     23     1.6    211  85  H55483       CHR220422 Homo sapien  6.12e-05
c    31     23     1.6    264  50  T83593       EST10843 Homo sapiens  6.12e-05
     32     23     1.6    306 107   HSC2GB051    H. sapiens partial cD  6.12e-05
     33     23     1.6    342  69  H74144       ys16b06.r1 Homo sapie  6.12e-05
c    34     23     1.6    344 159   AA034115     zi06d11.s1 Soares fet  6.12e-05
     35     23     1.6    425   9  T86331       yd94c12.r1 Homo sapie  6.12e-05
c    36     23     1.6    431  15  T11624       A1324F Homo sapiens c  6.12e-05
     37     23     1.6    441  18  T51819       yb54a10.s1 Homo sapie  6.12e-05
     38     23     1.6    464 165   D17509       Human placenta cDNA 5  6.12e-05
     39     23     1.6    485 173   W74443       zd75e11.s1 Soares fet  6.12e-05
c    40     23     1.6    566  46  R09449       yf19g10.s1 Homo sapie  6.12e-05
c    41     23     1.6    573 116   W26710       13e1 Human retina cDN  6.12e-05
     42     23     1.6    594 154   AA039629     mg95b11.r1 Soares mou  6.12e-05
     43     22     1.5    130  47  H45090       yo01g05.r1 Homo sapie  1.13e-03
     44     22     1.5    291 107   HSC0TA091    H. sapiens partial cD  1.13e-03
c    45     22     1.5    440 116   W27358       29g6 Human retina cDN  1.13e-03

                                 ALIGNMENTS

RESULT      1
LOCUS        H82541          454 bp    mRNA              EST        09-NOV-1995
DEFINITION   ys69h11.r1 Homo sapiens cDNA clone 220101 5' similar to
             gb:X51417_cds1 STEROID HORMONE RECEPTOR ERR2 (HUMAN);.
ACCESSION    H82542
NID          g1060631
KEYWORDS     EST.
```

```
SOURCE      human clone 220101 primer-M13RP1 library Soares retina N2b4HR
            vector-pT7T3D (Pharmacia) with a modified polylinker host=DH10B
            (ampicillin resistant) Rsite1=Not I Rsite2=Eco RI 1st strand cDNA
            was primed with a Not I - oligo(dT) primer
            [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCTTTTTTTTTTTTTTTTTT-3'],
            double-stranded cDNA was size selected, ligated to Eco RI adapters
            (Pharmacia), digested with Not I and cloned into the Not I and Eco
            RI sites of a modified pT7T3 vector (Pharmacia). The retinas were
            obtained from a 55 year old Caucasian male  and total cellular
            poly(A)+ RNA was extracted 6 hrs after their removal. The retina
            RNA was kindly provided by Roderick R. McInnes M.D. Ph.D. from the
            University of Toronto. Library constructed by Bento Soares and
            M.Fatima Bonaldo.
  ORGANISM  Homo sapiens
            Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
            Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
            Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
            Eutheria; Archonta; Primates, Catarrhini, Hominidae; Homo.
REFERENCE   1  (bases 1 to 454)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops. 345
            Source: IMAGE Consortium, LLNL
            This clone is available royalty free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
FEATURES             Location/Qualifiers
     source          1..454
                     /organism="Homo sapiens"
                     /clone="220101"
     mRNA            <1..>454
BASE COUNT       82 a    118 c    136 g    114 t      4 others
ORIGIN

  Query Match              3.7%;  Score 53;  DB 71;  Length 454;
  Best Local Similarity  65.7%;  Pred. No. 2.73e-56;
  Matches   119;  Conservative    0;  Mismatches  61;  Indels    1;  Gaps    1;

Db       137 cttcagcatccccactttgaggcatttcatgaagcggcaggcctggcaggacttgcgcct 196
             ||||| ||| |||||||  |     || |||     |   |||||||||||||||  ||||||||
Cp       489 CTTCACCATTCCCACTTCGTAACACTTCCGAAGTCGGCAGCCCTGGCAGCTCTTGCGCCC 430

Db       197 ccgtttggtgatctcgcactcgttggtggccgggcagctgtactcaatgttcccttggat 256
                 |||    |||     ||||  || || || || ||   | || ||| | |  ||||| ||
Cp       429 GTTTTTATCGATTGTACACTGATTTGTAGCTGGACAAATATAATCATTATGTCCTTGAAT 370

Db       257 agtcctcttgaagaaggccttgcaagcctcgcaggaggccaacgcctnagtggtagccag 316
               | || ||   | ||||||| ||   | || || ||   |   || ||   |||| || || |
Cp       369 GCTTCTTTTAAAAAAGGCCTTACATCCTTCACACGAC-CAGACTCCATAGTGATATCCCG 311

Db       317 a 317
             |
Cp       310 A 310


RESULT     2
LOCUS       W13715       390 bp    mRNA             EST       10-SEP-1996
DEFINITION  ma83b04.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone 317263 5'
            similar to gb:L13740 EARLY RESPONSE PROTEIN NAK1 (HUMAN);
ACCESSION   W13715
NID         g1287753
```

```
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 390)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: ETPrimer
            High quality sequence stop: 385.
FEATURES             Location/Qualifiers
     source          1..390
                     /organism="Mus musculus"
                     /note="Vector: pT7T3D (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCATTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library constructed by Bento
                     Soares and M.Fatima Bonaldo. RNA was kindly provided by
                     Dr. Minoru Ko (Wayne State University)."
                     /clone="317263"
                     /clone_lib="Soares mouse p3NMF19.5"
                     /dev_stage="19.5 dpc total fetus"
                     /lab_host="DH10B (ampicillin resistant)"
     mRNA            <1..>390
BASE COUNT       70 a    106 c    114 g    100 t
ORIGIN

  Query Match              3.6%;  Score 52;  DB 149;  Length 390;
  Best Local Similarity  62.7%;  Pred. No. 2.43e-54;
  Matches   128;  Conservative    0;  Mismatches   76;  Indels    0;  Gaps    0;

Db        80 cttaaccatcccgacactgagacacttctgaaatctgcagtactgacatcggtttcggcg 139
             ||| |||||| || ||     |   ||||||| ||| || ||||   ||| || |   || || ||
Cp       489 CTTCACCATTCCCACTTCGTAACACTTCCGAAGTCGGCAGCCCTGGCAGCTCTTGCGCCG 430

Db       140 tctcttgtccactgggcagttttatttgccaggcaaacatattttgcattttttctgcac 199
                 | || || | ||  || |  ||  | ||  | |||| ||| |     || |    || |
Cp       429 GTTTTTATCGATTGTACACTGATTTGTAGCTGGACAAATATAATCATTATGTCCTTGAAT 370

Db       200 cattctcttgaagaagccctgcagcctcccaggtccagactccatgatgctagcagaagc 259
                 |||| || || ||| ||| || || || || |   | |||||| ||||| |   |   |
Cp       369 GCTTCTTTTAAAAAAGGCCTTACATCCTTCACACGACCAGACTCCATAGTGATATCCCGA 310

Db       260 agcgttgtcgccgcacacggcaca 283
              ||||  ||||| ||| || || ||
Cp       309 TGCGTAATCGCTGCAGACAGGCCA 286


RESULT     3
LOCUS       H91890       403 bp    mRNA             EST       29-NOV-1995
DEFINITION  ys81d07.r1 Homo sapiens cDNA clone 221197 5' similar to
            gb:X51417_cds1 STEROID HORMONE RECEPTOR ERR2 (HUMAN);.
```

```
ACCESSION   H91890
NID         g1087468
KEYWORDS    EST.
SOURCE      human clone=221197 primer=M13RP1 library=Soares retina N2b4HR
            vector=pT7T3D (Pharmacia) with a modified polylinker host=DH10B
            (ampicillin resistant) Psite1=Not I Psite2=Eco RI 1st strand cDNA
            was primed with a Not I - oligo(dT) primer
            [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCTTTTTTTTTTTTTTTTTT-3'],
            double-stranded cDNA was size selected, ligated to Eco RI adapters
            (Pharmacia), digested with Not I and cloned into the Not I and Eco
            RI sites of a modified pT7T3 vector (Pharmacia). The retinas were
            obtained from a 55 year old Caucasian male  and total cellular
            poly(A)+ RNA was extracted 6 hrs after their removal. The retina
            RNA was kindly provided by Roderick R. McInnes M.D. Ph.D. from the
            University of Toronto. Library constructed by Bento Soares and
            M.Fatima Bonaldo.
  ORGANISM  Homo sapiens
            Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
            Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
            Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
            Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 403)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops: 330
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
FEATURES             Location/Qualifiers
     source          1..403
                     /organism="Homo sapiens"
                     /clone="221197"
     mRNA            <1..>403
BASE COUNT       73 a    113 c    118 g     97 t       2 others
ORIGIN

  Query Match              3.6%;  Score 52;  DB 74;  Length 403;
  Best Local Similarity  66.7%;  Pred. No. 2.43e-54;
  Matches   104;  Conservative    0;  Mismatches  52;  Indels    0;  Gaps    0;

Db     112 cttcagcatccccactttgaggcatttcatgaaggggcaggcctggcaggacttgcgcct 171
            ||||| ||| ||||||| |    || |||    |  ||||||||||||||  ||||||||
Cp     489 CTTCACCATTCCCACTTCCTAACACTTCCCAAGTCCGCAGGCCTGGCAGCTCCTGCGCCC 430

Db     172 ccgtttggtgatctcgcactcgttggtggccgggcagctgtactcaatgttcccttggat 231
               |||    |||     ||||  || || || || ||  | || ||| | |  ||||| ||
Cp     429 GTTTTTATCGATTGTACACTGATTTGTACCTGGACAAATATAATCATTATGTCCTTGAAT 370

Db     232 agtcttcttgaataaacccttacaaatccttgcaca 267
              | || || || |||||||| ||   | || || ||
Cp     369 GCTTCTTTTAAAAAAGGCCTTACATCCTTCACACGA 334


RESULT     4
LOCUS       AA157597      412 bp    mRNA             EST       11-DEC-1996
DEFINITION  zo44a02.r1 Stratagene endothelial cell 937223 Homo sapiens cDNA
            clone 589706 5' similar to gb:X07282 RETINOIC ACID RECEPTOR BETA-2
            (HUMAN);.
ACCESSION   AA157597
NID         g1729205
```

```
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 412)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E.,
            Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 351.
FEATURES             Location/Qualifiers
     source          1..412
                     /organism="Homo sapiens"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI; Cloned unidirectionally.  Primer: Oligo dT.
                     Umbilical vein endothelial cells, passaged once. Average
                     insert size: 1.0 kb; Uni-ZAP XR Vector;  5' adaptor
                     sequence: 5' GAATTCGGCACGAG 3' -3' adaptor sequence: 5'
                     CTCGAGTTTTTTTTTTTTTTTTTT 3'"
                     /clone="589706"
                     /clone_lib="Stratagene endothelial cell 937223"
                     /dev_stage="umbilical vein, 1 passage"
                     /lab_host="SOLR (kanamycin resistant)"
     mRNA            <1..>412
BASE COUNT      129 a     89 c    101 g     89 t       4 others
ORIGIN

  Query Match              3.6%;  Score 51;  DB 191;  Length 412;
  Best Local Similarity  64.9%;  Pred. No. 2.12e-52;
  Matches   111;  Conservative    0;  Mismatches  60;  Indels    0;  Gaps    0;

Db      29 gggtaccactatggggtcagcgcctgtgagggatgtaagggctttttccgcaggaatatt 88
            || || |||||||| ||| |   | ||||| |||||||||| ||||||     ||||| |||
Qy     313 GGATATCACTATGGAGTCTGGTCGTGTGAAGGATGTAAGGCCTTTTTTAAAACAACCATT 372

Db      89 cagaagaatatgatttacacttctcaccgagataagaactgtgttattaataaagtcacc 148
            ||      |||    ||| | |||||            ||  | |||    ||  ||||| |
Qy     373 CAAGGACATAATGATTATATTTGTCCAGCTACAAATCAGTGTACAATCGATAAAAACCCG 432

Db     149 aggaatcgatgccaatactgtcgactccagaagtgctttgaagtgggaatg 199
             | ||  | |||||   ||| ||||| | |||||| |  ||||||||||||
Qy     433 CCCAAGAGTTTCCAGGCTTGCCCATTTCCGAAGTGTTATGAAGTGGGAATG 483


RESULT     5
LOCUS       W53545        508 bp    mRNA             EST       03-JUN-1996
DEFINITION  md56a06.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
            clone 272372 5' similar to gb:X57528 M.musculus mRNA for retinoic
            acid receptor-alpha (MOUSE);.
ACCESSION   W53545
NID         g1357435
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
            Murinae; Mus.
REFERENCE   1  (bases 1 to 508)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:233802
            Seq primer: mob REGA+ET
            High quality sequence stop: 352.
FEATURES             Location/Qualifiers
     source          1..508
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGGAAATTTTTTTTTTTTTTTTTTTTTT
                     T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
                     14.5dpc embryos [total RNA provided by  Minoru Ko, Wayne
                     State Univ., from 2 ]; double-stranded cDNA was ligated to
                     Eco RI adaptors (Pharmacia), digested with Not I and
                     cloned into the Not I and Eco RI sites of the modified
                     pT7T3 vector. Library went through one round of
                     normalization, and was constructed by Bento Soares and
                     M.Fatima Bonaldo."
                     /clone="372370"
                     /clone_lib="Soares mouse embryo NbME13.5 14.5"
                     /sex="unknown"
                     /tissue_type="embryo"
                     /dev_stage="13.5-14.5dpc total fetus"
                     /lab_host="DH10B"
     mRNA            <1..>508
BASE COUNT      123 a    168 c    127 g     90 t
ORIGIN

  Query Match             3.3%;  Score 48;  DB 120;  Length 508;
  Best Local Similarity   63.6%;  Pred. No. 1.20e-46;
  Matches  112;  Conservative    0;  Mismatches   64;  Indels    0;  Gaps    0;

Db      233 catccggctaccactatggggtcagcgcctgtgagggctgtaaggacttcttccgacgaa 292
            |||| || || |||||||| ||| |  | ||||| || ||||||| ||| ||   |  ||
Qy      308 CATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATGTAAGGCCTTTTTTAAAAGAA 367

Db      293 gcatccagaagaacatggtgtatacgtgtcaccgggacaagaactgcatcatcaacaagg 352
            |||| ||     | |  | |||| ||||        ||  | ||  | ||| | ||
Qy      368 GCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCAGTGTACAATCGATAAAA 427

Db      353 tgacccggaaccgctgccagtactgccggctgcagaaatgtttcgacgtgggcatg 408
                 || ||  ||||||||  |||||| || | ||| |||| ||| ||||| |||
Qy      428 ACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTACGAAGTGGGAATG 483


RESULT      6
LOCUS       AA003987         536 bp    mRNA            EST       23-JUL-1996
DEFINITION  mq80b08.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
            clone 439287 5' similar to gb:X57528 M.musculus mRNA for retinoic
            acid receptor-alpha (MOUSE);.
ACCESSION   AA003987
NID         g1447462
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
            Murinae; Mus.
REFERENCE   1  (bases 1 to 536)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:264623
            Seq primer: mob.REGA+ET
            High quality sequence stop: 346.
FEATURES             Location/Qualifiers
     source          1..536
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGGAAATTTTTTTTTTTTTTTTTTTTTT
                     T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
                     14.5dpc embryos [total RNA provided by  Minoru Ko, Wayne
                     State Univ., from 2 ]; double-stranded cDNA was ligated to
                     Eco RI adaptors (Pharmacia), digested with Not I and
                     cloned into the Not I and Eco RI sites of the modified
                     pT7T3 vector. Library went through one round of
                     normalization, and was constructed by Bento Soares and
                     M.Fatima Bonaldo."
                     /clone="439287"
                     /clone_lib="Soares mouse embryo NbME13.5 14.5"
                     /sex="unknown"
                     /tissue_type="embryo"
                     /dev_stage="13.5-14.5dpc total fetus"
                     /lab_host="DH10B"
     mRNA            <1..>536
BASE COUNT      129 a    180 c    131 g     96 t
ORIGIN

  Query Match             3.3%;  Score 48;  DB 147;  Length 536;
  Best Local Similarity   63.6%;  Pred. No. 1.20e-46;
  Matches  112;  Conservative    0;  Mismatches   64;  Indels    0;  Gaps    0;

Db      286 catccggctaccactatggggtcagcgcctgtgagggctgtaaggacttcttccgacgaa 345
            |||| || || |||||||| ||| |  | ||||| || ||||||| ||| ||   |  ||
Qy      308 CATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATGTAAGGCCTTTTTTAAAAGAA 367

Db      346 gcatccagaagaacatggtgtatacgtgtcaccgggacaagaactgcatcatcaacaagg 405
            |||| ||     | |  | |||| ||||        ||  | ||  | ||| | ||
Qy      368 GCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCAGTGTACAATCGATAAAA 427

Db      406 tgacccggaaccgctgccagtactgccggctgcagaaatgtttcgacgtgggcatg 461
                 || ||  ||||||||  |||||| || | ||| |||| ||| ||||| |||
Qy      428 ACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTACGAAGTGGGAATG 483


RESULT      7
LOCUS       W53477           552 bp    mRNA            EST       03-JUN-1996
DEFINITION  md45d07.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
            clone 371341 5' similar to gb:X07282 RETINOIC ACID RECEPTOR BETA-2
            (HUMAN); gb:X57528 M.musculus mRNA for retinoic acid receptor-alpha
            (MOUSE);.
ACCESSION   W53477
```

```
NID         g1357321
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
            Murinae; Mus.
REFERENCE   1  (bases 1 to 552)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:232773
            Seq primer: mob.REGA+ET
            High quality sequence stop: 319.
FEATURES             Location/Qualifiers
     source          1..552
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI. 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGGAAATTTTTTTTTTTTTTTTTTTT
                     T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
                     14.5dpc embryos [total RNA provided by  Minoru Ko, Wayne
                     State Univ., from 2 ]; double-stranded cDNA was ligated to
                     Eco RI adaptors (Pharmacia), digested with Not I and
                     cloned into the Not I and Eco RI sites of the modified
                     pT7T3 vector. Library went through one round of
                     normalization, and was constructed by Bento Soares and
                     M.Fatima Bonaldo."
                     /clone="371341"
                     /clone_lib="Soares mouse embryo NbME13.5 14.5"
                     /sex="unknown"
                     /tissue_type="embryo"
                     /dev_stage="13.5-14.5dpc total fetus"
                     /lab_host="DH10B"
     mRNA            <1..>552
BASE COUNT      140 a    151 c    168 g     93 t
ORIGIN

  Query Match             2.9%;  Score 41;  DB 120;  Length 552;
  Best Local Similarity  62.2%;  Pred. No. 1.17e-33;
  Matches  130;  Conservative    0;  Mismatches  77;  Indels   2;  Gaps   2;

Db      230 ctgcttcgtttgccaggacaagtcatcgggctaccactatgggtgagtgcctgcaggag 289
            ||||    || |||    ||  |  |||||||| || ||||||| |||  |   | || || ||
Qy      285 CTGCCTGTCTGCAGCGATTACGCATCGGGATATCATTATGGAGTTTGGTCCTGTGAAGG 344

Db      290 gtgcaagggcttttccgcagaagtattca-gaagaacatgatctacacttgccatcgag 348
             || ||:| ||||::     ||||| ||||| | | |  |||||       |  |||  |  |
Qy      345 ATGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTAC 404

Db      349 ataggaactgcgtcattaacaaggt-actaggaa-cgatg-cagtactgccgcctgcaga 408
             | ||  | ||    ||  | ||   |     | ||  | ||||||  |||||| || | ||
Qy      405 A-AATCAGTGTACAATCGATAAAAATTGGTGAAGAGTTGCCAGGCTTGCCGACTTCGGA 463

Db      409 agtgctttgaagtgggcatgtcaaagagt 437
            |||| |  |||||||| |||   ||| ||
Qy      464 AGTGTTACGAAGTGGGAATGGTGAAGTGT 492
```

```
RESULT     8
LOCUS       AA092184       365 bp    mRNA            EST       24-OCT-1996
DEFINITION  [illegible] seq F Fetal heart, Lambda ZAP Express Homo sapiens cDNA 5'.
ACCESSION   AA092184
NID         g1637173
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 365)
  AUTHORS   Liew,C.C.
  TITLE     cDNAs from fetal heart (1996)
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Liew CC
            Molecular Cardiology
            University of Toronto
            Banting Institute, 100 College St., Toronto, Ontario, M5G1L5
            Tel: 4169788758
            Fax: 4169785650
            Email: liewcc@utcc.utoronto.ca
            PCR PRimers
            FORWARD: 5' GCCAAGCTCGAAATTAACCCTCACTAAAGGG 3'
            BACKWARD: 5' CCAGTGAATTGTAATACGACTCACTATAGGGCG 3'
            Seq primer: 5' GAAATTAACCCTCACTAAAGGG 3'.
FEATURES             Location/Qualifiers
     source          1..365
                     /organism="Homo sapiens"
                     /note="Vector: Lambda ZAP Express; Site_1: EcoRI; Site_2:
                     XhoI. mRNA was purified from human fetal hearts (8-10
                     weeks). cDNA was synthesized using a XhoI-Olido dT
                     adaptor-primer.  EcoRI adaptors were ligated, followed by
                     digestion with XhoI, for directional cloning into
                     predigested lambda ZAP Express."
                     /clone_lib="Fetal heart, Lambda ZAP Express"
                     /lab_host="E. coli XL1-Blue"
     mRNA            <1..>365
BASE COUNT       88 a     85 c    103 g     89 t
ORIGIN

  Query Match             2.6%;  Score 38;  DB 181;  Length 365;
  Best Local Similarity  77.1%;  Pred. No. 2.53e-28;
  Matches   54;  Conservative    0;  Mismatches  16;  Indels   0;  Gaps   0;

Db      164 ccttcagtgaagcttcgatgatgggcttactgaccaacctggcagacagggagctggtt 223
            ||||||  || || || ||||||   |  ||||||    |||| ||||     |  |  |
Qy      703 CCTTCACCGAGGCCTCCATGATGATCCCCTGACCAACTTCCCGACAAGGAGTTGGTA 762

Db      224 cacatgatca 233
            ||||||||||
Qy      763 CACATGATCA 772


RESULT     9
LOCUS       R52304         393 bp    mRNA            EST       18-MAY-1995
DEFINITION  yg[illegible] r1 Homo sapiens cDNA clone 39602 5' similar to gb:M29960
            STEROID RECEPTOR TR2 (HUMAN);.
ACCESSION   R52304
NID         g814206
KEYWORDS    EST.
SOURCE      human clone=39602 library=Soares infant brain 1NIB vector=Lafmid BA
            host=DH10B (ampicillin resistant) primer=M13RP1 Rsite1=Not I
            Rsite2=Hind III Whole brain from a 73 days post natal female. 1st
            strand cDNA was primed with a Not I - oligo(dT) primer [5'
            AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTTTT 3']; double-stranded
            cDNA was ligated to Hind III adaptors (Pharmacia), digested with
            Not I and directionally cloned into the Not I and Hind III sites of
            the lafmid BA vector. Library went through one round of
            normalization. Library constructed by Bento Soares and M.Fatima
            Bonaldo.
  ORGANISM  Homo sapiens
            Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
            Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
            Sarcopterygii; Choanata; Tetrapoda, Amniota, Mammalia, Theria,
            Eutheria; Archonta; Primates, Catarrhini, Hominidae, Homo
REFERENCE   1  (bases 1 to 393)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            GDB: G00-412-143
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops: 303
            Source: IMAGE Consortium, LLNL
            This clone is available royalty free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information
FEATURES             Location/Qualifiers
     source          1..393
                     /organism="Homo sapiens"
                     /clone="39602"
BASE COUNT      122 a     88 c     84 g     97 t      2 others
ORIGIN

  Query Match              2.6%;  Score 38;  DB 26;  Length 393;
  Best Local Similarity  66.7%;  Pred. No. 2.53e-28;
  Matches    74;  Conservative    0;  Mismatches  37;  Indels    0;  Gaps    0;

Db     271 aggaccaaataaggtttttgatctttgcgtagtatgtggagacaaagcatcaggacgtca 330
            |||  |||| | || |   | |   | ||||  || ||   | ||   | ||||| |||   |||
Qy     261 AGGTTCAAACAGGGATGCTCACTTCTGCGCTGTCTGCAGCGATTACGCATCGGATATCA 320

Db     331 ttatggagcagtaacttgtgaaggctgcaaaggtttttttaaaagnagcat 381
              |||||||       | |||||||| || || |  ||||||||||| |||||
Qy     321 CTATGGAGTCTGGTCCGTGTGAAGGATGTAAGGCCTTTTTTAAAAGAAGCAT 371


RESULT    10
LOCUS       AA049907      604 bp    mRNA              EST       30-DEC-1996
DEFINITION  mj13e03.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
            clone 475996 5' similar to gb:M73779 RETINOIC ACID RECEPTOR ALPHA-1
            (HUMAN); gb:X57528 M.musculus mRNA for retinoic acid receptor-alpha
            (MOUSE);.
ACCESSION   AA049907
NID         g1529596
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes, Metazoa, Chordata;
            Vertebrata, Eutheria; Rodentia, Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 604)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:286740
            Seq primer:  28M13 rev2 from Amersham
            High quality sequence stop: 359.
FEATURES             Location/Qualifiers
     source          1..604
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGGAAATTTTTTTTTTTTTTTTTTTTTT
                     T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
                     14.5dpc embryos [total RNA provided by  Minoru Ko, Wayne
                     State Univ., from 2 ]; double-stranded cDNA was ligated to
                     Eco RI adaptors (Pharmacia), digested with Not I and
                     cloned into the Not I and Eco RI sites of the modified
                     pT7T3 vector. Library went through one round of
                     normalization, and was constructed by Bento Soares and
                     M.Fatima Bonaldo."
                     /clone="475996"
                     /clone_lib="Soares mouse embryo NbME13.5 14.5"
                     /sex="unknown"
                     /tissue_type="embryo"
                     /dev_stage="13.5-14.5dpc total fetus"
                     /lab_host="DH10B"
     mRNA            <1..>604
BASE COUNT      142 a    194 c    153 g    115 t
ORIGIN

  Query Match              2.6%;  Score 38;  DB 195;  Length 604;
  Best Local Similarity  60.8%;  Pred. No. [illegible];
  Matches   107;  Conservative    0;  Mismatches  69;  Indels    0;  Gaps    0;

Db     287 catccggctaccactatgggtcagcgcctgtgagggctgtaagggcttcttccgacgaa 346
            |||| || || |||||||| ||| |   | ||||| || ||||||| ||| ||    | |||
Qy     308 CATCGGATATCACTATGGAGTCTGGTCCGTGTGAAGGATGTAAGGCCTTTTTTAAAAGAA 367

Db     347 gcatccagaagaacatgatgtatacgtgtcacccggacaagaactgcatcatcaacaagg 406
            |||| ||      | |  |  ||||  ||||            ||  | || |   ||| | ||
Qy     368 GCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCAGTGTACAATCGATAAAA 427

Db     407 tgacccggaacccgctgcagtactgccggctgcagaaatgtttcgacgttgggatg 462
                 || ||  |     |||  |||||| || | ||| |||| ||| || || |||
Qy     428 ACGGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTACGAAGTGGGAATG 483


RESULT    11
LOCUS       W53511      566 bp    mRNA              EST       03-JUN-1996
DEFINITION  md45f07.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
            clone 371365 5' similar to gb:X87283 RETINOIC ACID RECEPTOR BETA 2
            (HUMAN); gb:X57528 M.musculus mRNA for retinoic acid receptor-alpha
            (MOUSE);.
ACCESSION   W53511
NID         g1357336
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
            Murinae; Mus.
REFERENCE   1  (bases 1 to 566)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,I., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,R.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
```

```
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway  Box 8501  St. Louis, M  63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image llnl gov) for further information.
            MGI:232797
            Seq primer: mob PEGA-FT
            High quality sequence stop: 360.
FEATURES             Location/Qualifiers
     source          1..566
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /note="Vector  pT7T3D-Pac (Pharmacia) with a modified
                     polylinker: Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGGAAATTTTTTTTTTTTTTTTTTTTTT
                     T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
                     14.5dpc embryos [total RNA provided by  Minor, K.  Wayne
                     State Univ., from 2 ]; double-stranded cDNA was ligated to
                     Eco RI adaptors (Pharmacia), digested with Not I and
                     cloned into the Not I and Eco RI sites of the modified
                     pT7T3 vector  Library went through one round of
                     normalization, and was constructed by Bento Soares and
                     M.Fatima Bonaldo."
                     /clone="371365"
                     /clone_lib="Soares mouse embryo NbME13 5 14 5"
                     /sex="unknown"
                     /tissue_type="embryo"
                     /dev_stage="13.5-14.5dpc total fetus"
                     /lab_host="DH10B"
     mRNA            <1..>566
BASE COUNT      142 a    158 c    170 g     96 t
ORIGIN

  Query Match              2.5%,  Score 36,  DB 120,  Length 566,
  Best Local Similarity  62 7%;  Pred  No. 7 37e-25,
  Matches   126;  Conservative    0;  Mismatches  72,  Indels   3,  Gaps   3,

Db     242 ctgcttcgtttgccaggacaagtcatcgggctaccactatggcgtcagtgcctgcgaggg 301
            ||||   || |||    ||  |   ||||||| || |||||||| ||| |    | || || ||
Qy     285 CTGCGCTGTCTGCAGCGATTATGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGG 344

Db     302 [illegible] 360
             || |||| ||||||     ||||| ||||| | | |   |||||        |  ||| | |
Qy     345 ATGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTAC 404

Db     361 ataagaactgcgtcattaacaaggtcactaggaacgatgccagtactgccgccttgcag 420
            | ||   | ||    ||  | ||    |      | ||   | ||||||  |||||| ||| | |
Qy     405 A-AATCAGTGTACAATCGATAAAAACCGGCGTAAGACCTGCCAGGCCTGCCGATTT-CCC 462

Db     421 aagtgctttgaagtgggcatg 441
            |||||| |   ||||||||| |||
Qy     463 AAGTGTTACGAAGTGGGAATG 483


RESULT   12
LOCUS       C08482       378 bp    mRNA             EST       09-SEP-1996
DEFINITION  C.elegans cDNA clone yk77h11 . 5' end, single read.
ACCESSION   C08482
NID         g1533553
KEYWORDS    EST(expressed sequence tag).
SOURCE      Caenorhabditis elegans (strain:N2) embryo hermaphrodite cDNA to
            mRNA. clone_lib:Yuji Kohara unpublished cDNA clone:yk77h11.
  ORGANISM  Caenorhabditis elegans
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Nematoda;
            Secernentea; Rhabditia; Rhabditida; Rhabditina; Rhabditoidea;
            Rhabditidae; Caenorhabditis.
REFERENCE   1  (bases 1 to 378)
```

```
  AUTHORS   Kohara,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-APR-1996) to the DDBJ/EMBL/GenBank databases. Yuji
            Kohara, National Institute of Genetics, Gene Library Lab; Yata
            1111, Mishima, Shizuoka 411, Japan (E-mail:ykohara lab.nig.ac.jp,
            Tel:0559-81-6854, Fax:0559-81-6855)
REFERENCE   2  (sites)
  AUTHORS   Kohara,Y , Motohashi,T , Tabara,H , Watanabe,H., Sugimoto,A.,
            Sano,M., Miyata,A. and Nishigaki,A.
  TITLE     Expression map of the C.elegans genome
  JOURNAL   Unpublished (1996)
FEATURES             Location/Qualifiers
     source          1..378
                     /organism="Caenorhabditis elegans"
                     /strain="N2"
                     /clone="yk77h11"
                     /clone_lib="Yuji Kohara unpublished cDNA"
                     /dev_stage="embryo"
                     /sex="hermaphrodite"
BASE COUNT      126 a     75 c     89 g     82 t      6 others
ORIGIN

  Query Match              2.2%,  Score 32,  DB 141,  Length 378;
  Best Local Similarity  65 2%;  Pred  No. 3.44e-18;
  Matches    58,  Conservative    0,  Mismatches  31,  Indels    0,  Gaps    0;

Db     122 ttttgngcagtttgngggatacggcactgggaaagcactacggtgtaacggcttgtnat 181
            || || || || ||   | |||    |||   |||| | ||||| || ||    | | ||
Qy     283 TTCTGCGCTGTCTGCAGCGATTATGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAA 342

Db     182 ggatgcaagggatttnncaaaagaagtat 210
            ||||| ||||   |||    | |||||| ||
Qy     343 GGATGTAAGGCCTTTTTTAAAAGAAGCAT 371


RESULT   13
LOCUS       T83788       163 bp    mRNA             EST       16-MAR-1995
DEFINITION  yd65a11.r1 Homo sapiens cDNA clone 113084 5' similar to gb:M73779
            RETINOIC ACID RECEPTOR ALPHA-1 (HUMAN);.
ACCESSION   T83788
NID         g712076
KEYWORDS    EST.
SOURCE      human clone=113084 library=Soares fetal liver spleen 1NFLS
            vector=pT7T3D (Pharmacia) with a modified polylinker host=DH10B
            (ampicillin resistant) primer=M13RP1 Rsite1=Pac I Rsite2=Eco RI
            Liver and spleen from a 20 week-post conception male fetus. 1st
            strand cDNA was primed with a Pac I - oligo(dT) primer [5'
            AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTT 3'], double-stranded
            cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac
            I and cloned into the Pac I and Eco RI sites of the modified pT7T3
            vector  Library went through one round of normalization. Library
            constructed by Bento Soares and M.Fatima Bonaldo.
  ORGANISM  Homo sapiens
            Eukaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 163)
  AUTHORS   Hillier,L , Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M , Kucaba,T , Le,M , Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence starts: 1
            High quality sequence stops: 1
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality.
FEATURES             Location/Qualifiers
     source          1..163
                     /organism="Homo sapiens"
                     /clone="113084"
BASE COUNT       50 a     41 c     49 g     23 t
ORIGIN

  Query Match             2.0%;  Score 28;  DB 8;  Length 163;
  Best Local Similarity  78.4%;  Pred. No. 1.84e-13;
  Matches    40;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Db      35 cggaaccgctgccagtactgccgactgcagaagtgctttgaagtgggcatg 85
           || ||  ||||||| ||||||||| | |||||| |  |||||||| |||
Qy     433 CGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTACGAAGTGGGAATG 483


RESULT    14
LOCUS       CELK105GXF    360 bp    mRNA            EST       13-DEC-1995
DEFINITION  C.elegans cDNA clone yk105g10 : 5' end, single read.
ACCESSION   D75525
NID         g1121318
KEYWORDS    EST(expressed sequence tag).
SOURCE      Caenorhabditis elegans (strain N2, ) (library: Yuji kohara
            unpublished cDNA) Hermaphrodite embryo embryo cDNA to mRNA
  ORGANISM  Caenorhabditis elegans
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Nematoda;
            Secernentea; Rhabditia; Rhabditida; Rhabditina; Rhabditoidea;
            Rhabditidae; Caenorhabditis.
REFERENCE   1  (bases 1 to 360)
  AUTHORS   Kohara,Y., Motohashi,T., Tabara,H., Sugimoto,A., Watanabe,H. and
            Nishigaki,A.
  TITLE     Toward an expression map of the C.elegans genome
  JOURNAL   Unpublished (1995)
COMMENT     Submitted (23-Aug-1995) to DDBJ by:
            Yuji Kohara
            Gene Library Lab.
            National Institute of Genetics
            Yata 1111, Mishima Shizuoka
            411 Japan
            Phone: 0559-81-6854
            Fax  : 0559-81-6855
            Email:ykohara@ddbj.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1..360
                     /organism="Caenorhabditis elegans"
                     /strain="N2"
                     /dev_stage="embryo "
                     /sex="Hermaphrodite"
                     /tissue_type="embryo"
                     /clone_lib="Yuji kohara unpublished cDNA"
BASE COUNT       95 a     86 c    104 g     65 t     10 others
ORIGIN

  Query Match             2.0%;  Score 28;  DB 89;  Length 360;
  Best Local Similarity  65.5%;  Pred. No. 6.03e-12;
  Matches    57;  Conservative    0;  Mismatches   30;  Indels    0;  Gaps    0;

Db      98 tgcgctgtntgcggggacgtggctctcggaaaacactacggagtgaacgcttgcaacggg 157
           |||||||| |:| | ||     ||     ||| | ||||| |||||     | ||  | ||
Qy     286 TGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGA 345

Db     158 tgcaaaggatttt*taggagaagcatt 184
           || || |  |||||||  |||||||||
Qy     346 TGTAAGGCCTTTTTTAAAAGAAGCATT 372


RESULT    15
LOCUS       AA044343      431 bp    mRNA            EST       04-SEP-1996
DEFINITION  zk54a08.r1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone
```

```
            486614 5'.
ACCESSION   AA044343
NID         g1522199
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 431)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 358.
FEATURES             Location/Qualifiers
     source          1..431
                     /organism="Homo sapiens"
                     /note="Organ: uterus; Vector: pT7T3-Pac; Site_1: Not I;
                     Site_2: Eco RI; 1st strand cDNA was primed with a Not I -
                     oligo(dT) primer [5'
                     AACTGGAAGAATTCGCGGCCGCCTTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization.  Library
                     constructed by M. Fatima Bonaldo."
                     /clone="486614"
                     /clone_lib="Soares pregnant uterus NbHPU"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="DH10B"
     mRNA            <1..>431
BASE COUNT       75 a    136 c    143 g     71 t      6 others
ORIGIN

  Query Match             2.0%;  Score 28;  DB 157;  Length 431;
  Best Local Similarity  75.9%;  Pred. No. 6.03e-12;
  Matches    41;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Db     192 [illegible] 245
           ||| |||||||||||||||  ||||| |||| |  | |  | | |||||||| :|
Qy     621 GGACGGCCTGAGCCCCGAGCAGCTAGTGCTCACCCTCCTGGAGGCTGAGCCGCC 674


Search completed: Thu Dec 18 18:14:05 1997
Job time : 790 secs.
```

```
*******************************************************************

                    MPSRCH                              (TM)

*******************************************************************

          Release 2.1D John F. Collins, Biocomputing Research Unit.
          Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
                    Distribution rights by IntelliGenetics, Inc.

MPsrch_nn   n.a. - n.a. database search, using Smith-Waterman algorithm

Run on:          Thu Dec 18 18:57:03 1997; MasPar time 679.80 Seconds
                                           944.633 Million cell updates/sec
Tabular output not generated.

Title:           >US-08-826-361A-2
Description:     (1-1251) from US08826361A.seq
Perfect Score:   1251
N.A. Sequence:       1 ATGAATTACAGCATTCCCAG ... ACCTCAGGCATGCGAGGTGA 1251
    Comp:              TACTTAATGTCGTAAGGGTC     TGGAGTCCGTACGCTCCACT

Scoring table:   TABLE default
                 Gap 6

Nmatch   STD :   Dbase 0;  Query 0

Searched:        707517 seqs  256659390 bases x 2

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database:        EST-STS
                   1:EST1 2:EST2 3:EST3 4:EST4 5:EST5 6:EST6 7:EST7 8:EST8
                   9:EST9 10:EST10 11:EST11 12:EST12 13:EST13 14:EST14
                   15:EST15 16:EST16 17:EST17 18:EST18 19:EST19 20:EST20
                   21:EST21 22:EST22 23:EST23 24:EST24 25:EST25 26:EST26
                   27:EST27 28:EST28 29:EST29 30:EST30 31:EST31 32:EST32
                   33:EST33 34:EST34 35:EST35 36:EST36 37:EST37 38:EST38
                   39:EST39 40:EST40 41:EST41 42:EST42 43:EST43 44:EST44
                   45:EST45 46:EST46 47:EST47 48:EST48 49:EST49 50:EST50
                   51:EST51 52:EST52 53:EST53 54:EST54 55:EST55 56:EST56
                   57:EST57 58:EST58 59:EST59 60:EST60 61:EST61 62:EST62
                   63:EST63 64:EST64 65:EST65 66:EST66 67:EST67 68:EST68
                   69:EST69 70:EST70 71:EST71 72:EST72 73:EST73 74:EST74
                   75:EST75 76:EST76 77:EST77 78:EST78 79:EST79 80:EST80
                   81:EST81 82:EST82 83:EST83 84:EST84 85:EST85 86:EST86
                   87:EST87 88:EST88 89:EST89 90:EST90 91:EST91 92:EST92
                   93:EST93 94:EST94 95:EST95 96:EST96 97:EST97 98:EST98
                   99:EST99
Database:        EST-STS-TWO
                   100:EST100 101:EST101 102:EST102 103:EST103 104:EST104
                   105:EST105 106:EST106 107:EST107 108:EST108 109:EST109
                   110:EST110 111:EST111 112:EST112 113:EST113 114:EST114
                   115:EST115 116:EST116 117:EST117 118:EST118 119:EST119
                   120:EST120 121:EST121 122:EST122 123:EST123 124:EST124
                   125:EST125 126:EST126 127:EST127 128:EST128 129:EST129
                   130:EST130 131:EST131 132:EST132 133:EST133 134:EST134
                   135:EST135 136:EST136 137:EST137 138:EST138 139:EST139
                   140:EST140 141:EST141 142:EST142 143:EST143 144:EST144
                   145:EST145 146:EST146 147:EST147 148:EST148 149:EST149
                   150:EST150 151:EST151 152:EST152 153:EST153 154:EST154
                   155:EST155 156:EST156 157:EST157 158:EST158 159:EST159
                   160:EST160 161:EST161 162:EST162 163:EST163 164:EST164
                   165:EST165 166:EST166 167:EST167 168:EST168 169:EST169
                   170:EST170 171:EST171 172:EST172 173:EST173 174:EST174
                   175:EST175 176:EST176 177:EST177 178:EST178 179:EST179
                   180:EST180 181:EST181 182:EST182 183:EST183 184:EST184
                   185:EST185 186:EST186 187:EST187 188:EST188 189:EST189
                   190:EST190 191:EST191 192:EST192 193:EST193 194:EST194
                   195:EST195 196:EST196 197:EST197 198:EST198

Statistics:      Mean 11.181;  Variance 1.969;  scale 5.679

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                SUMMARIES
                      %
Result              Query
    No.   Score     Match Length DB  ID             Description          Pred. No.
  -----------------------------------------------------------------------------
c     1      52       4.2    390 149  W13715       ma83b04.r1 Soares mou  1.00e-54
c     2      52       4.2    403  74  H91890      ys81d07.r1 Homo sapie  1.00e-54
c     3      53       4.2    454  71  H82542      ys69h11.r1 Homo sapie  1.12e-56
      4      51       4.1    412 121  AA107537     zt44a02.r1 Stratagene  5.76e-53
      5      48       3.8    508 120  W53545       md56a06.r1 Soares mou  5.03e-47
      6      48       3.8    536 147  AA003987     mq80b08.r1 Soares mou  5.03e-47
      7      41       3.3    552 120  W53477       md45d07.r1 Soares mou  5.13e-34
      8      38       3.0    565 181  AA092184     114609.seq.F Fetal he  1.14e-28
      9      38       3.0    393  26  R52304      yg79c04.r1 Homo sapie  1.14e-28
     10      38       3.0    604 125  AA119907     mj10c03.r1 Soares mou  1.14e-28
     11      36       2.9    565 120  W55511       md45f07.r1 Soares mou  3.42e-25
     12      32       2.6    378 141  C08482       C.elegans cDNA clone   1.68e-18
     13      29       2.3    163   8  T83788      yd65a11.r1 Homo sapie  9.46e-14
     14      28       2.2    360  83  CELK105F3XF C.elegans cDNA clone   3.16e-12
     15      27       2.2    368 198  N44459       SW3ICA1324SK Brugia m  9.74e-11
     16      28       2.2    431 137  AA044343     zk54a08.r1 Soares pre  3.16e-12
     17      27       2.2    483   6  T78977      yd21e03.r1 Homo sapie  9.74e-11
     18      28       2.2    578 125  W74899       me72d01.r1 Soares mou  3.16e-12
     19      25       2.0    120 141  C09766       C.elegans cDNA clone   7.18e-08
     20      25       2.0    360  83  CELK082E1F  C.elegans cDNA clone   7.18e-08
     21      25       2.0    382 118  W39474       zc76d10.r1 Pancreatic  7.18e-08
     22      25       2.0    383 146  T26960       NIB1014D12R Infant br  7.18e-08
     23      24       1.9    141  65  H28316      yl60e10.r1 Homo sapie  1.69e-06
     24      24       1.9    336 154  AA011708     zi03d03.r1 Soares fet  1.69e-06
c    25      24       1.9    340 154  AA011660     zi03d03.s1 Soares fet  1.69e-06
     26      24       1.9    416  49  M78027      EST01614 Homo sapiens  1.69e-06
     27      24       1.9    466 154  AA009675     zi03f11.r1 Soares fet  1.69e-06
     28      24       1.9    487  63  H23042      ym51d05.r1 Homo sapie  1.69e-06
     29      24       1.9    536 151  W20940       mb96d08.r1 Soares mou  1.69e-06
     30      23       1.8    211  65  H15487      CHR220422 Homo sapien  3.59e-05
c    31      23       1.8    264  80  T30393      EST15843 Homo sapiens  3.59e-05
     32      23       1.8    3?? 1??  HSZZGB031    H. sapiens partial cD  3.59e-05
c    33      22       1.8    311  88  H66756      yr84b12.r1 Homo sapie  6.79e-04
     34      23       1.8    342  68  H74144      ys16b06.r1 Homo sapie  3.59e-05
c    35      23       1.8    344 158  AA034113     zi06d11.s1 Soares fet  3.59e-05
c    36      22       1.8    360  35  RICR1514A   Rice cDNA, partial se  6.79e-04
     37      23       1.8    425   9  T86331      yd84c12.r1 Homo sapie  3.59e-05
c    38      23       1.8    431  15  T11624      A1324F Homo sapiens c  3.59e-05
     39      23       1.8    441  18  T51819      yb54a10.s1 Homo sapie  3.59e-05
     40      23       1.8    464 165  C17508       Human placenta cDNA 5  3.59e-05
c    41      22       1.8    468 169  W49765       zc41f06.r1 Soares sen  6.79e-04
     42      22       1.8    47?  25  R5??41      yj59h12.s1 Homo sapie  6.79e-04
     43      23       1.8    485 173  W74443       zd75e11.s1 Soares fet  3.59e-05
c    44      23       1.8    575 116  W26710       13e1 Human retina cDN  3.59e-05
     45      23       1.8    594 154  AA008625     mq85b11.r1 Soares mou  3.59e-05

                                ALIGNMENTS

RESULT      1
LOCUS       W13715          390 bp    mRNA              EST       10-SEP-1996
DEFINITION  ma83b04.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone 317263 5'
            similar to gb:L13740 EARLY RESPONSE PROTEIN NAK1 (HUMAN);.
ACCESSION   W13715
NID         g1287753
KEYWORDS    EST.
```

```
SOURCE       house mouse.
  ORGANISM   Mus musculus
             Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
             Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
             Mus.
REFERENCE    1  (bases 1 to 390)
  AUTHORS    Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
             Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
             Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
             Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
             Waterston,R.
  TITLE      The WashU-HHMI Mouse EST Project
  JOURNAL    Unpublished (1996)
COMMENT
             Contact: Marra M/Mouse EST Project
             WashU-HHMI Mouse EST Project
             Washington University School of MedicineP
             4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
             Tel: 314 286 1800
             Fax: 314 286 1810
             Email: mouseest@watson.wustl.edu
             This clone is available royalty-free through LLNL ; contact the
             IMAGE Consortium (info@image.llnl.gov) for further information.
             Seq primer: ETPrimer
             High quality sequence stop: 385
FEATURES             Location/Qualifiers
     source          1..390
                     /organism="Mus musculus"
                     /note="Vector: pT7T3D (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCATTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library constructed by Bento
                     Soares and M.Fatima Bonaldo. RNA was kindly provided by
                     Dr. Minoru Ko (Wayne State University)."
                     /clone="317263"
                     /clone_lib="Soares mouse p3NMF19.5"
                     /dev_stage="19.5 dpc total fetus"
                     /lab_host="DH10B (ampicillin resistant)"
     mRNA            <1..>390
BASE COUNT        70 a    106 c    114 g    100 t
ORIGIN

  Query Match             4.2%;  Score 52;  DB 149;  Length 390;
  Best Local Similarity   62.7%;  Pred. No. 1.00e-54;
  Matches  128;  Conservative    0;  Mismatches   76;  Indels    0;  Gaps    0;

Db       80 cttaaccatcccgacactgagacacttctgaaatctgcagtactgacatcggtttcggcg 139
               ||| |||||  | ||     |   ||||||| ||| || ||||  ||| || |  || || ||
Cp      489 CTTCACCATTCCCACTTCGTAACACTTCCGAAGTCGGCAGGCCTGGCAGCTCTTGCGCCG 430

Db      140 tctcttgtccactgggcagttttatttgccaggcaaacatattttgcatttttctgcac 199
                  |  | || | ||  || |  ||  | ||  | |||| ||| |    || |   || |
Cp      429 GTTTTATCGATTGTACACTGATTTGTAGCTGGACAAATATAATCATTATGTCCTTGAAT 370

Db      200 cgttctcttgaagaagcccttgcagccctcgcaggtgcggactccgtagtgctggcaggc 259
                  |||| || || ||| |||| || || || || |  | |||||| ||||| |  |  |
Cp      369 GCTTCTTTTAAAAAAGGCCTTACATCCTTCACACGACCAGACTCCATAGTGATATCCCGA 310

Db      260 agcgttgtcgccgcacacggcaca 283
               ||||  |||| ||| || || ||
Cp      309 TGCGTAATCGCTGCAGACAGCGCA 286


RESULT      2
LOCUS       H91890          403 bp    mRNA             EST       29-NOV-1995
DEFINITION  ys81d07.r1 Homo sapiens cDNA clone 221197 5' similar to
            gb:X51417_cds1 STEROID HORMONE RECEPTOR ERR2 (HUMAN);.
ACCESSION   H91890
```

```
NID          g1087468
KEYWORDS     EST.
SOURCE       human clone 221197 primer=M13RP1 library=Soares retina N2b4HR
             vector=pT7T3D (Pharmacia) with a modified polylinker host=DH10B
             (ampicillin resistant) Rsite1=Not I Rsite2=Eco RI 1st strand cDNA
             was primed with a Not I - oligo(dT) primer
             [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCTTTTTTTTTTTTTTTTTT-3'],
             double-stranded cDNA was size selected, ligated to Eco RI adapters
             (Pharmacia), digested with Not I and cloned into the Not I and Eco
             RI sites of a modified pT7T3 vector (Pharmacia). The retinas were
             obtained from a 55 year old Caucasian male  and total cellular
             poly(A)+ RNA was extracted 6 hrs after their removal. The retina
             RNA was kindly provided by Roderick R. McInnes M.D. Ph.D. from the
             University of Toronto. Library constructed by Bento Soares and
             M.Fatima Bonaldo.
  ORGANISM   Homo sapiens
             Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
             Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
             Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
             Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 403)
  AUTHORS    Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
             Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
             Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
             Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
             Wilson,R.
  TITLE      The WashU-Merck EST Project
  JOURNAL    Unpublished (1995)
COMMENT
             Contact: Wilson RK
             WashU-Merck EST Project
             Washington University School of Medicine
             4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
             Tel: 314 286 1800
             Fax: 314 286 1810
             Email: est@watson.wustl.edu
             High quality sequence stops: 330
             Source: IMAGE Consortium, LLNL
             This clone is available royalty free through LLNL ; contact the
             IMAGE Consortium (info@image.llnl.gov) for further information.
FEATURES             Location/Qualifiers
     source          1..403
                     /organism="Homo sapiens"
                     /clone="221197"
     mRNA            <1..>403
BASE COUNT       73 a    113 c    118 g     97 t      2 others
ORIGIN

  Query Match             4.2%;  Score 52;  DB 74;  Length 403;
  Best Local Similarity   66.7%;  Pred. No. 1.00e-54;
  Matches  104;  Conservative    0;  Mismatches   52;  Indels    0;  Gaps    0;

Db      112 cttcagcatccccactttgaggcatttcatgaagcggcaggcctggcaggacttgcgcct 171
            ||||| ||| |||||||  |    || |||    |  |||||.|||||||||  |||.||||
Cp      489 CTTCACCATTCCCACTTCGTAACACTTCCGAAGTCGGCAGGCCTGGCAGCTCTTGCGCCG 430

Db      172 ccgtttggtgatctcgcactcgttggtggccgggcagctgtactcaatgttcccttggat 231
               ||.    |||     ||||  || || || || ||   | || |||  | |   ||||| ||
Cp      429 GTTTTATCGATTGTACACTGATTTGTAGCTGGACAAATATAATCATTATGTCCTTGAAT 370

Db      232 aatcctcttgaagaaggccttgcaagcctcgcagga 267
               | || || || ||||||| ||   | || || ||
Cp      369 GCTTCTTTTAAAAAAGGCCTTACATCCTTCACACGA 334


RESULT      3
LOCUS       H82542          454 bp    mRNA             EST       09-NOV-1995
DEFINITION  ys69h11.r1 Homo sapiens cDNA clone 220101 5' similar to
            gb:X51417_cds1 STEROID HORMONE RECEPTOR ERR2 (HUMAN);.
ACCESSION   H82542
NID         g1060631
KEYWORDS    EST.
SOURCE      human clone-220101 primer-M13RP1 library-Soares retina N2b4HR
```

```
            vector-pT7T3D (Pharmacia) with a modified polylinker host=DH10B
            (ampicillin resistant) Rsite1-Not I Rsite2-Eco RI 1st strand cDNA
            was primed with a Not I - oligo(dT) primer
            [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCTTTTTTTTTTTTTTTTTT-3'],
            double-stranded cDNA was size selected, ligated to Eco RI adapters
            (Pharmacia), digested with Not I and cloned into the Not I and Eco
            RI sites of a modified pT7T3 vector (Pharmacia). The retinas were
            obtained from a 55 year old Caucasian male  and total cellular
            poly(A)+ RNA was extracted 6 hrs after their removal. The retina
            RNA was kindly provided by Roderick R  McInnes M.D. Ph.D. from the
            University of Toronto  Library constructed by Bento Soares and
            M.Fatima Bonaldo.
  ORGANISM  Homo sapiens
            Eukaryotae; Metazoa; Eumetazoa, Bilateria, Coelomata,
            Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
            Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
            Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 454)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K , Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L , Rohlfing,T , Soares,M , Tan,F ,
            Trevaskis,E , Waterston,R , Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St  Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops: 345
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL , contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
FEATURES             Location/Qualifiers
     source          1..454
                     /organism="Homo sapiens"
                     /clone="220101"
     mRNA            <1..>454
BASE COUNT       82 a    118 c    136 g    114 t      4 others
ORIGIN

  Query Match             4.2%;  Score 53;  DB 71;  Length 454;
  Best Local Similarity  65.7%;  Pred. No. 1.12e-56;
  Matches  119;  Conservative    0;  Mismatches  61;  Indels   1;  Gaps   1;

Db    137 cttcagcatccccactttgaggcatttcatgaagcggcaggcctggcaggacttgcgcct 196
          ||||| ||| |||||||| |    || |||    |  ||||||||||||||  ||||||||
Cp    489 CTTCACCATTCCCATTTTGTAATATTTTTGAAGTTGGTAGGCCTGGCAGTTTTTTTGCCG 430

Db    197 ccgtttggtgatctcgcactcgttggtggccgggcagctgtactcaatgttcccttggat 256
             |||   |||    ||||  || || || || || ||   | || ||| | |  ||||| ||
Cp    429 TTTTTTATCGATTGTATATTGATTTGTAGCTGGATAAATATAATCATTATGTCCTTGAAT 370

Db    257 agtcctcttgaagaaggccttgcaagcctcgcaggaggccaacgcctnagtggtagccag 316
            | || || || |||||||| ||   | || || ||   |  || ||   |||| || || |
Cp    369 GCTTCTTTTAAAAAAGGCCTTACATCCTTCACACGAC-CACACTCCATAGTGATATCCCG 311

Db    317 a 317
          |
Cp    310 A 310


RESULT      4
LOCUS       AA157597      412 bp    mRNA            EST       11-DEC-1996
DEFINITION  zo44a02.r1 Stratagene endothelial cell 937223 Homo sapiens cDNA
            clone 589706 5' similar to gb X07282 RETINOIC ACID RECEPTOR BETA-2
            (HUMAN);.
ACCESSION   AA157597
NID         g1729205
```

```
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 412)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E.,
            Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL , contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 351.
FEATURES             Location/Qualifiers
     source          1..412
                     /organism="Homo sapiens"
                     /note="Vector: pBluescript SK , Site_1: EcoRI; Site_2:
                     XhoI, Cloned unidirectionally.  Primer: Oligo dT.
                     Umbilical vein endothelial cells, passaged once. Average
                     insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor
                     sequence: 5' GAATTCGGCACGAG 3' -3' adaptor sequence: 5'
                     CTCGAGTTTTTTTTTTTTTTTTTT 3'"
                     /clone="589706"
                     /clone_lib="Stratagene endothelial cell 937223"
                     /dev_stage="umbilical vein, 1 passage"
                     /lab_host="SOLR (kanamycin resistant)"
     mRNA            <1..>412
BASE COUNT      129 a     89 c    101 g     89 t      4 others
ORIGIN

  Query Match             4.1%;  Score 51;  DB 191;  Length 412;
  Best Local Similarity  64.9%;  Pred. No. 8.73e-53;
  Matches  111;  Conservative    0;  Mismatches  60;  Indels    0;  Gaps    0;

Db     29 gggtaccactatggggtcagcgcctgtgagggatgtaagggctttttccgcaaagtatt 88
          || || |||||||| ||| |   | ||||| |||||||||| ||||||    ||||| |||
Qy    313 GGATATCATTATGGAGTTGGTTGTGTGAAGGATGTAAGGCTTTTTTTAAAAGAACCATT 372

Db     89 cagaagaatatgatttacacttgtcaccgagataagaactgtgttattaataaagtcacc 148
          ||     |||     ||| | |||||           ||   | |||    ||   ||||   |
Qy    373 CAAGGACATAATGATTATATTTGTTTAGTTATAAATTAGTGTATAATTGATAAAAAACGG 432

Db    149 aggaatcgatgccaatactgtcgactccagaagtgctttgaagtgggaatg 199
            | ||   | |||||    ||| ||||| | |||||| |   |||||||||||
Qy    433 CGCAAGAAGTTGTCAGGCCTACCGACTTCGGAAGTGTTACGAAGTGGGAATG 483


RESULT      5
LOCUS       W53545        508 bp    mRNA            EST       03-JUN-1996
DEFINITION  md56a06.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
            clone 372370 5' similar to gb:X57528 M.musculus mRNA for retinoic
            acid receptor-alpha (MOUSE);.
ACCESSION   W53545
NID         g1357435
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
            Murinae; Mus.
REFERENCE   1  (bases 1 to 508)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:233802
            Seq primer: mob.REGA+ET
            High quality sequence stop: 352.
FEATURES             Location/Qualifiers
     source          1..508
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGGAAATTTTTTTTTTTTTTTTTTTTT
                     T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
                     14.5dpc embryos [total RNA provided by  Minoru Ko, Wayne
                     State Univ., from 2 ]; double-stranded cDNA was ligated to
                     Eco RI adaptors (Pharmacia), digested with Not I and
                     cloned into the Not I and Eco RI sites of the modified
                     pT7T3 vector. Library went through one round of
                     normalization, and was constructed by Bento Soares and
                     M.Fatima Bonaldo."
                     /clone="372370"
                     /clone_lib="Soares mouse embryo NbME13.5 14.5"
                     /sex="unknown"
                     /tissue_type="embryo"
                     /dev_stage="13.5-14.5dpc total fetus"
                     /lab_host="DH10B"
     mRNA            <1..>508
BASE COUNT      123 a    168 c    127 g     90 t
ORIGIN

  Query Match             3.8%;  Score 48;  DB 120;  Length 508;
  Best Local Similarity   63.6%;  Pred. No. 5.03e-47;
  Matches  112;  Conservative    0;  Mismatches   64;  Indels    0;  Gaps    0;

Db        233 catccggctaccactatggggtcagcgcctgtgagggctgtaagggcttcttccgacgaa 292
              |||| || || |||||||| ||| |   | ||||| || ||||||| ||| ||    | |||
Qy        308 CATGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATGTAAGGCCTTTTTTTAAAAGAA 367

Db        293 gcatccagaaggacatggtgtatacgtgtcaccgggacaaaaactgcatcatcaacaagg 352
              |||| ||     | |  |  ||||  ||||          ||  | || |  ||| | ||
Qy        368 GCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCAGTGTACAATCGATAAAA 427

Db        353 tgacccggaaccgctgccagtactgccggctgcagaaatgtttcgacgtgggcatg 408
                  || ||  ||||||||  |||||| || | ||| |||| ||| ||||| |||
Qy        428 ACGGCGGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTACGAAGTGGGAATG 483


RESULT      6
LOCUS       AA003987       536 bp    mRNA            EST       23-JUL-1996
DEFINITION  mq80b08.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
            clone 439287 5' similar to gb:X57528 M.musculus mRNA for retinoic
            acid receptor-alpha (MOUSE);.
ACCESSION   AA003987
NID         g1447462
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
            Murinae; Mus.
REFERENCE   1  (bases 1 to 536)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:264623
            Seq primer: mob.REGA+ET
            High quality sequence stop: 346.
FEATURES             Location/Qualifiers
     source          1..536
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGGAAATTTTTTTTTTTTTTTTTTTTT
                     T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
                     14.5dpc embryos [total RNA provided by  Minoru Ko, Wayne
                     State Univ., from 2 ]; double-stranded cDNA was ligated to
                     Eco RI adaptors (Pharmacia), digested with Not I and
                     cloned into the Not I and Eco RI sites of the modified
                     pT7T3 vector. Library went through one round of
                     normalization, and was constructed by Bento Soares and
                     M.Fatima Bonaldo."
                     /clone="439287"
                     /clone_lib="Soares mouse embryo NbME13.5 14.5"
                     /sex="unknown"
                     /tissue_type="embryo"
                     /dev_stage="13.5-14.5dpc total fetus"
                     /lab_host="DH10B"
     mRNA            <1..>536
BASE COUNT      129 a    180 c    131 g     96 t
ORIGIN

  Query Match             3.8%;  Score 48;  DB 147;  Length 536;
  Best Local Similarity   63.6%;  Pred. No. 5.03e-47;
  Matches  112;  Conservative    0;  Mismatches   64;  Indels    0;  Gaps    0;

Db        286 catccggctaccactatggggtcagcgcctgtgagggctgtaagggcttcttccgacgaa 345
              |||| || || |||||||| ||| |   | ||||| || ||||||| ||| ||    | |||
Qy        308 CATGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATGTAAGGCCTTTTTTTAAAAGAA 367

Db        346 gcatccagaaggacatggtgtatacgtgtcaccgggacaaaaactgcatcatcaacaagg 405
              |||| ||     | |  |  ||||  ||||          ||  | || |  ||| | ||
Qy        368 GCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCAGTGTACAATCGATAAAA 427

Db        406 tgacccggaaccgctgccagtactgccggctgcagaaatgtttcgacgtgggcatg 461
                  || ||  ||||||||  |||||| || | ||| |||| ||| ||||| |||
Qy        428 ACGGCGGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTACGAAGTGGGAATG 483


RESULT      7
LOCUS       W53477         552 bp    mRNA            EST       03-JUN-1996
DEFINITION  md45d07.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
            clone 371341 5' similar to gb:X07282 RETINOIC ACID RECEPTOR BETA-2
            (HUMAN); gb:X57528 M.musculus mRNA for retinoic acid receptor-alpha
            (MOUSE);.
ACCESSION   W53477
```

```
NID         g1357321
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
            Murinae; Mus.
REFERENCE   1  (bases 1 to 552)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:232773
            Seq primer: mob REGA+ET
            High quality sequence stop: 319.
FEATURES             Location/Qualifiers
     source          1..552
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGCAAATTTTTTTTTTTTTTTTTTTTTTT
                     T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
                     14.5dpc embryos [total RNA provided by  Minoru Ko, Wayne
                     State Univ., from 2 ]; double-stranded cDNA was ligated to
                     Eco RI adaptors (Pharmacia), digested with Not I and
                     cloned into the Not I and Eco RI sites of the modified
                     pT7T3 vector. Library went through one round of
                     normalization, and was constructed by Bento Soares and
                     M.Fatima Bonaldo."
                     /clone="371341"
                     /clone_lib="Soares mouse embryo NbME13.5 14.5"
                     /sex="unknown"
                     /tissue_type="embryo"
                     /dev_stage="13.5-14.5dpc total fetus"
                     /lab_host="DH10B"
     mRNA            <1..>552
BASE COUNT      140 a    151 c    168 g     93 t
ORIGIN

  Query Match             3.3%;  Score 41;  DB 120;  Length 552;
  Best Local Similarity  62.2%;  Pred. No. 5.13e-34;
  Matches  130;  Conservative    0;  Mismatches  77;  Indels    2;  Gaps    2;

Db     230 [illegible] 289
               ||||    || |||    |   |   ||||||| || |||||||| |||  |    |  || || ||
Qy     285 CTGCGCTGTCTGCAGGGATTACGCATCGGGATATCACTATGGAGTTTGGTCGGTGTCAAGC 344

Db     290 gtgcaagggcttttccgcagaagtattca-gaagaacatgatctacacttgccatcgag 348
                || |||| ||||||     ||||| ||||| | | |  |||||        |  |||| | |
Qy     345 ATGTAAGGCCTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTAC 404

Db     349 ataagaactgcgtcattaacaggtcactaggaccgatgccagtactgccgcctgcaga 408
               | ||   | ||     || | ||    |     | ||  | ||||||  |||||| || | ||
Qy     405 A-AATCAGTGTACAATCGATAAAAACCAGCGTAAGAGCTGCTAGCCTGCCGACTTCGGA 463

Db     409 agtgctttgaagtgggcatgtcaaagagt 437
               |||| |   ||||||||| |||   ||| ||
Qy     464 AGTGTTACGAAGTGGGAATGGTGAAGTGT 492
```

```
RESULT     8
LOCUS       AA092184      365 bp    mRNA             EST       24-OCT-1996
DEFINITION  114609.seq.F Fetal heart, Lambda ZAP Express Homo sapiens cDNA 5'.
ACCESSION   AA092184
NID         g1637173
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 365)
  AUTHORS   Liew,C.C.
  TITLE     cDNAs from fetal heart (1996)
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Liew CC
            Molecular Cardiology
            University of Toronto
            Banting Institute, 100 College St., Toronto, Ontario, M5G1L5
            Tel: 4169788758
            Fax: 4169785650
            Email: liewcc@utcc.utoronto.ca
            PCR PRimers
            FORWARD: 5' GCCAAGCTCGAAATTAACCCTCACTAAAGGG 3'
            BACKWARD: 5' CCAGTGAATTGTAATACGACTCACTATAGGGCG 3'
            Seq primer: 5' GAAATTAACCCTCACTAAAGGG 3'.
FEATURES             Location/Qualifiers
     source          1..365
                     /organism="Homo sapiens"
                     /note="Vector: Lambda ZAP Express; Site_1: EcoRI; Site_2:
                     XhoI; mRNA was purified from human fetal hearts (8-10
                     weeks).  cDNA was synthesized using a XhoI-Oligo dT
                     adaptor-primer.  EcoRI adaptors were ligated  followed by
                     digestion with XhoI, for directional cloning into
                     predigested lambda ZAP Express."
                     /clone_lib="Fetal heart, Lambda ZAP Express"
                     /lab_host="E. coli XL1-Blue"
     mRNA            <1..>365
BASE COUNT       88 a     85 c    103 g     89 t
ORIGIN

  Query Match             3.0%;  Score 38;  DB 181;  Length 365;
  Best Local Similarity  77.1%;  Pred. No. 1.14e-28;
  Matches   54;  Conservative    0;  Mismatches  16;  Indels    0;  Gaps    0;

Db     164 [illegible] 223
               |||||||  || || || ||||||    |  ||||||||  |||| |||| |||| ||||
Qy     703 [illegible] 762

Db     224 cacatgatca 233
               ||||||||||
Qy     763 CACATGATCA 772


RESULT     9
LOCUS       R52304        393 bp    mRNA             EST       18-MAY-1995
DEFINITION  yg74e04.r1 Homo sapiens cDNA clone 39602 5' similar to gb:M29960
            STEROID RECEPTOR TR2 (HUMAN);.
ACCESSION   R52304
NID         g814206
KEYWORDS    EST.
SOURCE      human clone=39602 library=Soares infant brain 1NIB vector=Lafmid BA
            host=DH10B (ampicillin resistant) primer=M13RP1 Rsite1=Not I
            Rsite2=Hind III Whole brain from a 73 days post natal female. 1st
            strand cDNA was primed with a Not I - oligo(dT) primer [5'
            AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTTTT 3'], double-stranded
            cDNA was ligated to Hind III adaptors (Pharmacia), digested with
            Not I and directionally cloned into the Not I and Hind III sites of
            the Lafmid BA vector. Library went through one round of
            normalization. Library constructed by Bento Soares and M.Fatima
            Bonaldo.
  ORGANISM  Homo sapiens
            Eukaryotae, Metazoa, Eumetazoa, Bilateria; Coelomata;
            Deuterostomia, Chordata; Vertebrata; Gnathostomata; Osteichthyes;
            Sarcopterygii; Choanata; Tetrapoda; Amniota, Mammalia, Theria,
            Eutheria, Archonta, Primates, Catarrhini, Hominidae, Homo
REFERENCE   1  (bases 1 to 393)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            GDB: G00-412-143
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops. 303
            Source: IMAGE Consortium, LLNL
            This clone is available royalty free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information
FEATURES             Location/Qualifiers
     source          1..393
                     /organism="Homo sapiens"
                     /clone="39602"
BASE COUNT      122 a     88 c     84 g     97 t      2 others
ORIGIN

  Query Match             3.0%;  Score 38;  DB 26;  Length 393;
  Best Local Similarity  66.7%;  Pred. No. 1.14e-28;
  Matches   74;  Conservative    0;  Mismatches  37;  Indels    0;  Gaps    0;

Db     271 aggaccaaataaggttttgatctttgcgtagtatgtggagacaaagcatcaggacgtca 330
           |||  |||| | || |  | |  | ||||  || ||  | ||  | ||||| |||  |||
Qy     261 AGGTTCAAAGAGGGATGCTCACTTCTGCGCTGTCTGCAGCGATTACGCATCGGGATATCA 320

Db     331 ttatggagcagtaacttgtgaaggctgcaaaggtttttttaaaagnagcat 381
            |||||||      | |||||||| || || |  ||||||||||| |||||
Qy     321 CTATGGAGTCTGGTCGTGTGAAGGATGTAAGGCCTTTTTTAAAAGAAGCAT 371


RESULT    10
LOCUS       AA049907       604 bp    mRNA             EST       30-DEC-1996
DEFINITION  mj13e03.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
            clone 475996 5' similar to gb:M73779 RETINOIC ACID RECEPTOR ALPHA-1
            (HUMAN); gb:X57528 M.musculus mRNA for retinoic acid receptor alpha
            (MOUSE);.
ACCESSION   AA049907
NID         g1529596
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae, mitochondrial eukaryotes, Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 604)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
```

```
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:286740
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 359.
FEATURES             Location/Qualifiers
     source          1..604
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGGAAATTTTTTTTTTTTTTTTTTTTT
                     T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
                     14.5dpc embryos [total RNA provided by  Minoru Ko, Wayne
                     State Univ., from 2 ], double-stranded cDNA was ligated to
                     Eco RI adaptors (Pharmacia), digested with Not I and
                     cloned into the Not I and Eco RI sites of the modified
                     pT7T3 vector. Library went through one round of
                     normalization, and was constructed by Bento Soares and
                     M.Fatima Bonaldo."
                     /clone="475996"
                     /clone_lib="Soares mouse embryo NbME13.5 14.5"
                     /sex="unknown"
                     /tissue_type="embryo"
                     /dev_stage="13.5-14.5dpc total fetus"
                     /lab_host="DH10B"
     mRNA            <1..>604
BASE COUNT      142 a    194 c    153 g    115 t
ORIGIN

  Query Match             3.0%;  Score 38;  DB 195;  Length 604;
  Best Local Similarity  60.8%;  Pred. No. 1.14e-28;
  Matches  107;  Conservative    0;  Mismatches  69;  Indels    0;  Gaps    0;

Db     287 catccggctaccactatggggtcagcgcctgtgagggctgtaagggcttcttccgacgaa 346
           |||| || || |||||||| ||| |  | ||||| || ||||||| ||| ||   | |||
Qy     308 CATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATGTAAGGCCTTTTTTAAAAGAA 367

Db     347 gcatccagaagaacatggtgtatacgtgtcaccgggacaagaactgcatcatcaacaaag 406
           |||| ||     | |  |  ||||  ||||        ||  | || |  ||| | |||
Qy     368 GCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCAGTGTACAATCGATAAAA 427

Db     407 taacctggaaccggctgcagtactgccggctgcagaaatgtttcgacgttgggatg 462
                 | ||  ||   |||  |||||| || | ||| |||| ||| || || |||
Qy     428 ACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAACTGTTACGAAGTGGGAATG 483


RESULT    11
LOCUS       W53511         566 bp    mRNA             EST       03-JUN-1996
DEFINITION  md45f07.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
            clone 371365 5' similar to gb:X07282 RETINOIC ACID RECEPTOR BETA 2
            (HUMAN); gb:X57528 M.musculus mRNA for retinoic acid receptor-alpha
            (MOUSE);.
ACCESSION   W53511
NID         g1357336
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae, mitochondrial eukaryotes, Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
            Murinae; Mus.
REFERENCE   1  (bases 1 to 566)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
```

```
COMMENT
            Contact. Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway  Box 8501  St  Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:232797
            Seq primer: mob.REGA+ET
            High quality sequence stop: 360.
FEATURES             Location/Qualifiers
     source          1..566
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker  Site_1  Not I  Site_2  Eco RI. 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGAGCGGCCGCGGAAATTTTTTTTTTTTTTTTTTTTT
                     T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
                     14.5dpc embryos [total RNA provided by  Minoru Ko, Wayne
                     State Univ.  from 2 ]; double-stranded cDNA was ligated to
                     Eco RI adaptors (Pharmacia)  digested with Not I and
                     cloned into the Not I and Eco RI sites of the modified
                     pT7T3 vector  Library went through one round of
                     normalization, and was constructed by Bento Soares and
                     M.Fatima Bonaldo."
                     /clone="371365"
                     /clone_lib="Soares mouse embryo NbME13.5 14.5"
                     /sex="unknown"
                     /tissue_type="embryo"
                     /dev_stage="13 5-14 5dpc total fetus"
                     /lab_host="DH10B"
     mRNA            <1..>566
BASE COUNT      142 a    158 c    170 g     96 t
ORIGIN

  Query Match             2.9%,  Score 36,  DB 120,  Length 566,
  Best Local Similarity  62.7%;  Pred  No  3 42e-25;
  Matches   126;  Conservative    0;  Mismatches  72;  Indels    3;  Gaps    3;

Db    242 [illegible] 301

Qy    287 [illegible] 344

Db    302 [illegible] 360

Qy    345 ATGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTAC 404

Db    361 [illegible] 420

Qy    405 A-AATCAGTGTACAATCGATAAAAACCGGCGTAAGAGCTGCCAGGCCTGCCGACTT-CGG 462

Db    421 aagtgctttgaagtgggcatg 441

Qy    463 AAGTGTTACGAAGTGGGAATG 483


RESULT   12
LOCUS       C08482          378 bp    mRNA            EST       09-SEP-1996
DEFINITION  C.elegans cDNA clone yk77h11 : 5' end, single read
ACCESSION   C08482
NID         g1533553
KEYWORDS    EST(expressed sequence tag).
SOURCE      Caenorhabditis elegans (strain N2) embryo hermaphrodite cDNA to
            mRNA  clone_lib:Yuji Kohara unpublished cDNA clone:yk77h11
  ORGANISM  Caenorhabditis elegans
            Eukaryotae  mitochondrial eukaryotes  Metazoa  Nematoda
            Secernentea: Rhabditia: Rhabditida: Rhabditina; Rhabditoidea;
            Rhabditidae; Caenorhabditis.
REFERENCE   1  (bases 1 to 378)
  AUTHORS   Kohara,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-APR-1996) to the DDBJ/EMBL/GenBank databases. Yuji
            Kohara, National Institute of Genetics, Gene Library Lab; Yata
            1111, Mishima, Shizuoka 411, Japan (E-mail:ykohara@lab.nig.ac.jp,
            Tel:0559-81-6854, Fax:0559-81-6855)
REFERENCE   2  (sites)
  AUTHORS   Kohara,Y., Motohashi,T., Tabara,H., Watanabe,H., Sugimoto,A.,
            Sano,M., Miyata,A. and Nishigaki,A.
  TITLE     Expression map of the C.elegans genome
  JOURNAL   Unpublished (1996)
FEATURES             Location/Qualifiers
     source          1..378
                     /organism="Caenorhabditis elegans"
                     /strain="N2"
                     /clone="yk77h11"
                     /clone_lib="Yuji Kohara unpublished cDNA"
                     /dev_stage="embryo"
                     /sex="hermaphrodite"
BASE COUNT      126 a     75 c     89 g     82 t      6 others
ORIGIN

  Query Match             2.6%;  Score 32;  DB 141;  Length 378;
  Best Local Similarity  65.2%;  Pred. No. 1.68e-18;
  Matches    58;  Conservative    0;  Mismatches  31;  Indels    0;  Gaps    0;

Db    122 [illegible] 181

Qy    293 [illegible] 342

Db    182 [illegible] 210

Qy    343 GGATGTAAGGCCTTTTTTAAAACAAGCAT 371


RESULT   13
LOCUS       T83788          163 bp    mRNA            EST       16-MAR-1995
DEFINITION  yd65a11.r1 Homo sapiens cDNA clone 113084 5' similar to gb:M73779
            RETINOIC ACID RECEPTOR ALPHA-1 (HUMAN);.
ACCESSION   T83788
NID         g712076
KEYWORDS    EST.
SOURCE      human clone=113084 library-Soares fetal liver spleen 1NFLS
            vector=pT7T3 (Pharmacia) with a modified polylinker host=DH10B
            (ampicillin resistant) primer=M13RP1 Rsite1=Pac I Rsite2=Eco RI
            Liver and spleen from a 20 week-post conception male fetus. 1st
            strand cDNA was primed with a Pac I - oligo(dT) primer [5'
            AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTT 3'], double-stranded
            cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac
            I and cloned into the Pac I and Eco RI sites of the modified pT7T3
            vector.  Library went through one round of normalization. Library
            constructed by Bento Soares and M.Fatima Bonaldo.
  ORGANISM  Homo sapiens
            Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 163)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence starts: 1
            High quality sequence stops: 1
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality.
FEATURES             Location/Qualifiers
     source          1..163
                     /organism="Homo sapiens"
                     /clone="113084"
BASE COUNT       50 a     41 c     49 g     23 t
ORIGIN

  Query Match             2.3%;  Score 29;  DB 8;  Length 163;
  Best Local Similarity  78.4%;  Pred. No. 9.46e-14;
  Matches   40;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Db       35 cggaaccgctgccagtactgccgactgcagaagtgctttgaagtgggcatg 85
            || ||  ||||||||  ||||||||| | |||||| |  |||||||| |||
Qy      433 CGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTACGAAGTGGGAATG 483


RESULT    14
LOCUS       CELK105GXF     360 bp    mRNA            EST       13-DEC-1995
DEFINITION  C.elegans cDNA clone yk105g10 : 5' end, single read
ACCESSION   D75525
NID         g1121318
KEYWORDS    EST(expressed sequence tag).
SOURCE      Caenorhabditis elegans (strain N2 ) (library: Yuji kohara
            unpublished cDNA) Hermaphrodite embryo embryo cDNA to mRNA
  ORGANISM  Caenorhabditis elegans
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Nematoda;
            Secernentea; Rhabditia; Rhabditida; Rhabditina; Rhabditoidea;
            Rhabditidae; Caenorhabditis.
REFERENCE   1  (bases 1 to 360)
  AUTHORS   Kohara,Y., Motohashi,T., Tabara,H., Sugimoto,A., Watanabe,H. and
            Nishigaki,A.
  TITLE     Toward an expression map of the C.elegans genome
  JOURNAL   Unpublished (1995)
COMMENT     Submitted (23-Aug-1995) to DDBJ by:
            Yuji Kohara
            Gene Library Lab.
            National Institute of Genetics
            Yata 1111, Mishima Shizuoka
            411 Japan
            Phone: 0559-81-6854
            Fax  : 0559-81-6855
            Email:ykohara@ddbj.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1..360
                     /organism="Caenorhabditis elegans"
                     /strain="N2"
                     /dev_stage="embryo"
                     /sex="Hermaphrodite"
                     /tissue_type="embryo"
                     /clone_lib="Yuji kohara unpublished cDNA"
BASE COUNT       95 a     86 c    104 g     65 t     10 others
ORIGIN

  Query Match             2.2%;  Score 28;  DB 89;  Length 360;
  Best Local Similarity  65.5%;  Pred. No. 3.16e-12;
  Matches   57;  Conservative    0;  Mismatches   30;  Indels    0;  Gaps    0;

Db       98 tgcgctgtntgcggggacgtggctctcggaaaacactacggagtgaacgcttgcaacggg 157
            ||||||||| ||| | ||      ||     ||| | ||||| |||||     | ||  | ||
Qy      286 TCGCTGTCTGCAGCGATTACGCATCGGATATCACTATGGAGTCTGGTCGTGTGAAGGA 345

Db      158 tgcaaaggattttttaggagaagcatt 184
            || || |   |||||||  |||||||||
Qy      346 TCTAAGGCCTTTTTTAAAAGAAGCATT 372


RESULT    15
LOCUS       N44459          368 bp    mRNA            EST       12-DEC-1996
DEFINITION  SW3ICA1324SK Brugia malayi infective larva cDNA (SAW94WL-BmL3)
            Brugia malayi cDNA clone SW3ICA1324 5'.
ACCESSION   N44459
NID         g1182978
KEYWORDS    EST.
SOURCE      Brugia malayi.
  ORGANISM  Brugia malayi
            Eukaryotae; mitochondrial eukaryotes, Metazoa, Nematoda;
            Secernentea; Spiruria; Spirurida; Spirurina; Filarioidea;
            Onchocercidae; Brugia.
REFERENCE   1  (bases 1 to 368)
  AUTHORS   Blaxter,M.L., Raghavan,N., Ghosh,I., Guiliano,D., Lu,W.,
            Williams,S.A., Slatko,B. and Scott,A.L.
  TITLE     Genes expressed in Brugia malayi infective third stage larvae
  JOURNAL   Mol. Biochem. Parasitol. 77 (1), 77-93 (1996)
COMMENT
            Contact: Steven A. Williams
            Molecular Parasitology
            Smith College Department of Biological Sciences
            Department of Biological Sciences, Clark Science Center, Smith
            College, Northampton, MA, 01063, USA
            Tel: 4135853826
            Fax: 4135853786
            Email: genome@smith edu
            Seq primer: pBluescript pBluescript.
FEATURES             Location/Qualifiers
     source          1..368
                     /organism="Brugia malayi"
                     /strain="TRS Labs"
                     /note="Vector: lambda UniZap XR; Site_1: EcoR I; Site_2:
                     Xho I; Lymphatic filarial nematode parasite of humans.
                     mRNA was prepared from third stage infective larvae of
                     Brugia malayi isolated from mosquitoes 10 days after
                     infection and converted to double stranded cDNA using
                     reverse transcriptase and oligo(dT) followed by RNase H
                     and DNApol I. The library had 1 6 x 10E6 independent
                     recombinants and average insert size was 900 base pairs.
                     The library was constructed by Wenhong Lu. The library is
                     available from Dr. S.A. Williams, email genome@smith edu."
                     /clone="SW3ICA1324"
                     /clone_lib="Brugia malayi infective larva cDNA
                     (SAW94WL-BmL3)"
                     /lab_host="XL1-Blue MRF'"
     mRNA            <1..>368
BASE COUNT      105 a     51 c     79 g    109 t     24 others
ORIGIN

  Query Match             2.2%;  Score 27;  DB 198;  Length 368;
  Best Local Similarity  51.9%;  Pred. No. 9.74e-11;
  Matches  107;  Conservative    0;  Mismatches   99;  Indels    0;  Gaps    0;

Db      162 ctgatttctgtgtcggntgtggtgataaggntatnggtaagcattatggtactgtatcat 221
            || | ||||| |   |  ||   | ||| | |       ||   | || ||||| |       | |
Qy      278 CTCACTTCTGGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGT 337

Db      222 gtaatggcatgtaaaggttnctcnggcaggaaggaggtaagaaagattngnattatactngtc 281
            || | || ||||| |  |  |      ||||||          |  |     ||| || | |||
Qy      338 GTGAAGGATGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATACTTGTC 397

Db      282 gnttcagnaaacaatgcaacattgataaagatcatcggaatgcatgccgttatnatcgcn 341
                 | || || || |  || |||||| | |  || ||     ||||        | ||
Qy      398 CAGCTACAAATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGAC 457

Db      342 tccagaaatgtttngnagatggnatg 367
            | | ||| ||||  | ||  || |||
Qy      458 TTCGGAAGTGTTACGAAGTGGGAATG 483

Search completed. Thu Dec 18 19.08.37 1997
Job time : 694 secs.
```

```
***********************************************************************
   __       __    _______   ______    ______    ______    _     _
  | |\    /| |   |   __  | |  ____|  |   _  |  |  ____|  | |   | |
  | | \  / | |   |  |__| | | |____   |  |_| |  | |       | |___| |
  | |  \/  | |   |  _____| |____  |  |  _  _|  | |       |  ___  |
  | |      | |   | |            | |  | | \ \   | |____   | |   | |
  |_|      |_|   |_|       |______|  |_|  \_\  |______|  |_|   |_|  (TM)

***********************************************************************

          Release 2 ID John F  Collins  Biocomputing Research Unit
          Copyright (c) 1993, 1994  1995  University of Edinburgh  U K
                   Distribution rights by IntelliGenetics  Inc

MPsrch_nn   n.a. -  n.a. database search, using Smith-Waterman algorithm

Run on:            Thu Dec 18 19:08:59 1997:  MasPar time 254 03 Seconds
                                              971 065 Million cell updates/sec
Tabular output not generated

Title:             >US-08-826-361A-2
Description:       (1-1251) from US08826361A seq
Perfect Score:     1251
N.A. Sequence          1 ATGAATTACAGCATTCCCAG ...     ... ACCTCAGGCATGCGAGGTGA 1251
     Comp.               TACTTAATGTCGTAAGGGTC ...     ... TGGAGTCCGTACGCTCCACT

Scoring table:     TABLE default
                   Gap 6

Nmatch   STD :     Dbase 0:  Query 0

Searched:          359085 seqs  137405154 bases x 2

Post-processing: Minimum Match 0%
                   Listing first 45 summaries

Database:          EST-STS-THREE
                     1:EST199 2:EST200 3:EST201 4:EST202 5:EST203 6:EST204
                     7:EST205 8:EST206 9:EST207 10:EST208 11:EST209 12:EST210
                     13:EST211 14:EST212 15:EST213 16:EST214 17:EST215
                     18:EST216 19:EST217 20:EST218 21:EST219 22:EST220
                     23:EST221 24:EST222 25:EST223 26:EST224 27:EST225
                     28:EST226 29:EST227 30:EST228 31:EST229 32:EST230
                     33:EST231 34:EST232 35:EST233 36:STS1 37:STS2 38:STS3
                     39:STS4 40:STS5 41:STS6 42:STS7 43:STS8 44:STS9 45:STS10
                     46:STS11 47:STS12 48:STS13
Database:          EST-STS-FOUR
                     49:gnEST1 50:gnEST2 51:gnEST3 52:gnEST4 53:gnEST5
                     54:gnEST6 55:gnEST7 56:gnEST8 57:gnEST9 58:gnEST10
                     59:gnEST11 60:gnEST12 61:gnEST13 62:gnEST14 63:gnEST15
                     64:gnEST16 65:gnEST17 66:gnEST18 67:gnEST19 68:gnEST20
                     69:gnEST21 70:gnEST22 71:gnEST23 72:gnEST24 73:gnEST25
                     74:gnSTS 75:enEST1 76:enEST2 77:enEST3 78:enEST4
                     79:enEST5 80:enEST6 81:enEST7 82:enEST8 83:enEST9
                     84:enEST10 85:enEST11 86:enEST12 87:enEST13 88:enEST14
                     89:enEST15 90:enEST16 91:enEST17 92:enEST18 93:enEST19
                     94:enEST20 95:enEST21 96:enEST22 97:enEST23 98:enEST24
                     99:enEST25 100:enEST26 101:enEST27 102:enEST28
                     103:enEST29 104:enEST30 105:enEST31 106:enSTS 107:ueEST1
                     108:ueEST2

Statistics:        Mean 11.207;  Variance 2.021;  scale 5.546

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                 SUMMARIES
                     %
Result             Query
```

```
       No.  Score  Match Length DB  ID           Description            Pred. No.
   -------------------------------------------------------------------------------
  c     1     52    4.2    390 96  MM7156       ma83b04.r1 Soares mou  1.49e-53
        2     52    4.2   2222 44  G28564       human STS SHGC-35195.  1.95e-55
  c     3     40    3.2    593 28  AA023625     mh80b06.r1 Soares mou  1.39e-31
        4     26    2.1    280 25  N53519       yv57g05.r1 Soares fet  3.29e-09
        5     24    1.9    156 36  DM36D7S      D. melanogaster STS d  1.72e-06
        6     24    1.9    534 37  MM94003      mb96d08.r1 Soares mou  1.72e-06
  c     7     22    1.8    326 85  HS1164487    111371.seq.F Human fe  5.94e-04
  c     8     22    1.8    326 69  AA249798     111371.seq.F Human fe  5.94e-04
        9     22    1.8    382 30  AA057786     zf15e04.s1 Soares fet  5.94e-04
  c    10     22    1.8    399 21  AA021604     ze67c04.s1 Soares ret  5.94e-04
       11     22    1.8    424 89  HSW0986      ze08h09.r1 Soares fet  5.94e-04
       12     22    1.8    424  2  W90098       ze08h09.r1 Soares fet  5.94e-04
       13     22    1.8    452 28  AA024507     ze75d08.s1 Soares fet  5.94e-04
  c    14     23    1.9    490 16  AA196299     zp95c10.s1 Stratagene  3.38e-05
  c    15     22    1.8    498 11  AA179181     zp46b09.r1 Stratagene  5.94e-04
  c    16     22    1.8    501 83  HS1151180    zr69e09.r1 Soares NhH  5.94e-04
  c    17     22    1.8    501 65  AA233250     zr69e09.r1 Soares NhH  5.94e-04
  c    18     22    1.8    545 18  AA202284     LD02433.5prime LD Dro  5.94e-04
       19     21    1.7    295 2   SS284156     S.scrofa mRNA, expres  9.26e-03
  c    20     21    1.7    314 44  G27335       human STS CHLC.ATA7CB  9.26e-03
       21     21    1.7    320  2  W95560       ze02b04.r1 Soares fet  9.26e-03
       22     21    1.7    320 90  HSW5605      ze02b04.r1 Soares fet  9.26e-03
  c    23     21    1.7    359 36  G00807       fruit fly STS Dm0700   9.26e-03
       24     21    1.7    364 66  AA275237     zs58g06.r1 Soares NhH  9.26e-03
       25     21    1.7    364 83  HS1153545    zs58g06.r1 Soares NhH  9.26e-03
  c    26     21    1.7    387 15  AA192801     zp93g06.s1 Stratagene  9.26e-03
       27     21    1.7    422  7  AA126160     zm78b03.r1 Stratagene  9.26e-03
       28     21    1.7    426 65  AA233020     zr46a08.r1 Soares NhH  9.26e-03
       29     21    1.7    426 93  HS1150940    zr46a08.r1 Soares NhH  9.26e-03
       30     21    1.7    438 43  G24207       human STS WI-15034.    9.26e-03
  c    31     21    1.7    445 69  AA226701     nc27g02.r1 NCI CGAP F  9.26e-03
  c    32     21    1.7    445 81  HS1146141    nc27g02.r1 NCI CGAP F  9.26e-03
  c    33     21    1.7    495 55  AA096657     mc03h09.r1 Stratagene  9.26e-03
  c    34     21    1.7    538 108  MM1139113    mv68c05.r1 Soares mou  9.26e-03
  c    35     21    1.7    538 49  AA220605     mv68c05.r1 Soares mou  9.26e-03
  c    36     21    1.7    546 22  AA211579     zn56b11.s1 Stratagene  9.26e-03
  c    37     21    1.7    565 69  AA208970     mw72g12.r1 Soares mou  9.26e-03
  c    38     21    1.7    565 103  MMAA81339    mw72g12.r1 Soares mou  9.26e-03
  c    39     21    1.7    565 26  AA208970     mw72g12.r1 Soares mou  9.26e-03
  c    40     21    1.7    576 19  AA205532     zq67a04.s1 Stratagene  9.26e-03
       41     21    1.7    595  7  AA123429     mp91a08.r1 Soares 2NL  9.26e-03
       42     21    1.7    596 95  MMAA23429    mp91a08.r1 Soares 2NL  9.26e-03
  c    43     21    1.7    617 35  AA111111     mx96b05.r1 Barstead m  9.26e-03
  c    44     21    1.7    667 17  AA144507     mt54g08.r1 Stratagene  9.26e-03
  c    45     21    1.7    667 51  AA138017     mt54g08.r1 Stratagene  9.26e-03

                               ALIGNMENTS

RESULT     1
ID   MM7156      standard, RNA, EST, 390 BP.
AC   W13715;
NI   g1287753
DI   30-APR-1996 (Rel. 47, Created)
DI   04-MAR-1997 (Rel. 51, Last updated, Version 2)
DE   ma83b04.r1 Soares mouse p3NMF19 5 Mus musculus cDNA clone 317263 5'
DE   similar to gb:L13740 EARLY RESPONSE PROTEIN NAK1 (HUMAN);.
KW   EST.
OS   Mus musculus (house mouse)
OC   Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
OC   Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
OC   Mus.
RN   [1]
RP   1-390
RA   Marra M., Hillier L., Allen M., Bowles M., Dietrich N.,
RA   Dubuque T., Geisel S., Kucaba T., Lacy M., Le M., Martin J.,
RA   Morris M., Schellenberg K., Steptoe M., Tan F., Underwood K.,
RA   Moore B., Theising B., Wylie T., Lennon G., Soares B., Wilson R.,
RA   Waterston R.;
RT   "The WashU-HHMI Mouse EST Project";
RL   Unpublished.
```

```
CC   Contact: Marra M/Mouse EST Project WashU-HHMI Mouse EST Project
CC   Washington University School of MedicineP 4444 Forest Park Parkway.
CC   Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810
CC   Email: mouseest@watson.wustl.edu This clone is available
CC   royalty-free through LLNL ; contact the IMAGE Consortium
CC   (info@image.llnl.gov) for further information. Seq primer: ETPrimer
CC   High quality sequence stop: 385.
FH   Key             Location/Qualifiers
FH
FT   source          1..390
FT                   /organism="Mus musculus"
FT                   /note="Vector: pT7T3D (Pharmacia) with a modified
FT                   polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
FT                   was primed with a Not I - oligo(dT) primer [5'
FT                   TGTTACCAATCTGAAGTGGGAGCGGCCGCATTTTTTTTTTTTTTTTTT 3'],
FT                   double-stranded cDNA was size selected, ligated to Eco RI
FT                   adapters (Pharmacia), digested with Not I and cloned into
FT                   the Not I and Eco RI sites of a modified pT7T3 vector
FT                   (Pharmacia). Library went through one round of
FT                   normalization to a Cot = 5. Library constructed by Bento
FT                   Soares and M Fatima Bonaldo. RNA was kindly provided by Dr

FT                   Minoru Ko (Wayne State University)."
FT                   /clone="317263"
FT                   /clone_lib="Soares mouse p3NMF19.5"
FT                   /dev_stage="19.5 dpc total fetus"
FT                   /lab_host="DH10B (ampicillin resistant)"
FT   mRNA            <1..>390
SQ   Sequence 390 BP; 70 A; 106 C; 114 G; 100 T; 0 other;

  Query Match             4.2%;  Score 52;  DB 96;  Length 390;
  Best Local Similarity   62.7%;  Pred. No. 1.49e-53;
  Matches  128;  Conservative    0;  Mismatches   76;  Indels    0;  Gaps    0;

Db       80 cttaaccatcccgacactgagacacttctgaaatctgcagtactgacatcggtttcggcg 139
            ||| ||||| || ||   |  ||||||| ||| || ||||  ||| || |  || || ||
Qy      489 CTTCACCATTCCCACTTCGTAACACTTCCGAAGTCGGCAGCCCTCCCAGCTCTTGCGCCG 430

Db      140 tctcttgtccactgagcagtttttatttgccaggcaaacatatttgcattttttctgcac 199
               | || || | ||  || |  ||   | ||   | |||| ||| |     || |   || |
Qy      429 GTTTTTATCGATTGTACACTGATTTGTAGCTGGACAAATATAATCATTATGTCCTTGAAT 370

Db      200 cgttctcttgaagaagcccttgcagccctcgcaggtgcggactccgtagtgctggcaggc 259
               |||| || || ||| |||| || || || || |   | |||||| ||||| |   |  |
Qy      369 GCTTCTTTTAAAAAAGGCCTTACATCCTTCACACGACCAGACTCCATAGTGATATCCCGA 310

Db      260 agcgttgtcgccgcacacggcaca 283
             ||||  |||| ||| || || ||
Qy      309 TGCGTAATCGCTGCAGACAGCGCA 286


RESULT      2
LOCUS       G28564        2202 bp    DNA             STS       01-JUL-1996
DEFINITION  human STS SHGC-35295.
ACCESSION   G28564
NID         g1408379
KEYWORDS    STS sequence; primer; sequence tagged site.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo
REFERENCE   1  (bases 1 to 2202)
  AUTHORS   Myers,R.M.
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Richard M. Myers
            Stanford Human Genome Center (SHGC)
            Stanford University School of Medicine
            Department of Genetics, M-344, Stanford, CA 94305, USA
            Tel: 4157259687
            Fax: 4157259689
            Email: myers@shgc.stanford.edu
```

```
            Primer A: AATTGAACTCACTGCTTTTGAGG
            Primer B: CAGCTTCAGGTATCATCTTCACC
            STS size: 206
            PCR Profile:
                  Initial incubation: 94 degrees C for 90 seconds

                  Denaturation:       94 degrees C for 15 seconds
                  Annealing:          62 degrees C for 23 seconds
                  Polymerization:     72 degrees C for 30 seconds
                  PCR Cycles:         30
                  Thermal Cycler:     Perkin Elmer 9600
            Protocol:
                  Template:        25 ng
                  Primer:          each 1 uM
                  dNTPs:           each 200 uM
                  Taq Polymerase:  0.05 units/ul
                  Total Vol:       10 ul

            Buffer:
                  MgCl2:           2.5 mM
                  KCl:             50  mM
                  Tris-HCl:        20  mM
                  pH:              8.3

            Prepared with primer pairs provided by Sandoz, derived from M29960
            - Washington University/Merck EST sequence.
FEATURES             Location/Qualifiers
     source          1..2202
                     /organism="Homo sapiens"
                     /map="12"
     STS             1926..2131
     primer_bind     1926..1948
     primer_bind     complement(2109..2131)
BASE COUNT      720 a    442 c    439 g    601 t
ORIGIN

  Query Match             4.2%;  Score 53;  DB 44;  Length 2202;
  Best Local Similarity   61.9%;  Pred. No. 1.85e-55;
  Matches  138;  Conservative    0;  Mismatches   85;  Indels    0;  Gaps    0;

Db      368 aggaccaaataaggtttttgatctttgcgtagtatgtggagacaaagcatcaggacgtca 427
             |||  |||| | || |   | |   | ||||  || ||   | ||   | ||||| |||  |||
Qy      261 AGGTTCAAAGACGGATGCTCACTTCTGCCCTGTCTCCACCCATTACCCATCCCGATATCA 320

Db      428 ttatggagcagtaacttgtgaaggctgcaaaggattttttaaaagaagcatccgaaaaaa 487
               |||||||       | |||||||| || || |  ||||||||||||||||| | |   | |
Qy      321 CTATGGAGTCTGGTCGTGTGAAGGATGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACA 380

Db      488 tttagtatattcatgtcgaggatcaaagattgtattattaataacacaccaccaaaccg 547
             |    |   |||    |||| ||    |||| | ||||  ||   ||||  |||  || ||  |
Qy      381 TAATGATTATATTGTCCAGCTACAAATCAGTGTACAATCGATAAAAACGGCCGCCAAGAG 440

Db      548 ctgtcaatactgcaggttacagaaatgtattgcgtttggaatg 590
             ||| ||   |||| |  | | ||  |||   |   | ||||||
Qy      441 CTGCCAGGCCTGCCGACTTCGGAAGTGTTACGAAGTGGGAATG 483


RESULT      3
LOCUS       AA023625       593 bp    mRNA            EST       21-JAN-1997
DEFINITION  mh80b05.r1 Soares mouse placenta 4NbMP13.5 14.5 Mus musculus cDNA
            clone 457235 5' similar to gb:M12674 ESTROGEN RECEPTOR (HUMAN);
            gb:M38651 Mouse estrogen receptor mRNA, complete cds (MOUSE);.
ACCESSION   AA023625
NID         g1487542
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 593)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway  Box 8501  St  Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL , contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:274123
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 492.
FEATURES             Location/Qualifiers
     source          1..593
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker: Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGGAAATTTTTTTTTTTTTTTTTTTTTT
                     T 3'], double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization, and was
                     constructed by Bento Soares and M.Fatima Bonaldo."
                     /clone="457235"
                     /clone_lib="Soares mouse placenta 4NbMP13.5 14.5"
                     /sex="unknown"
                     /dev_stage="adult"
                     /lab_host="DH10B"
     mRNA            1..593
BASE COUNT      144 a    165 c    119 g    165 t
ORIGIN

  Query Match             3.2%;  Score 40;  DB 29;  Length 593;
  Best Local Similarity   74.4%;  Pred. No. 1.39e-31;
  Matches   61;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Db        174 tctttgcccagttgatcatatgaaccagctccctatctgctaggttggtcaataagccca 233
              |||| |||||| ||||||| || |||| |||| | || || |  |||||||    |    ||
Qy        784 TCTTGGCCCAGCTGATCATGTGTACCAACTCCTTGTCGCCCAACTTGGTCAGGGACATCA 725

Db        234 tcattgaggcttcactgaaggg 255
              |||| |||||| ||  |||||||
Qy        724 TCATGGAGGCCTCGGTGAAGGG 703


RESULT     4
LOCUS       N59515         280 bp    mRNA            EST       28-JAN-1997
DEFINITION  yv57g05.r1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone
            246872 5' similar to SW.LRH1_MOUSE P45448 LIVER RECEPTOR HOMOLOG ;.
ACCESSION   N59515
NID         g1203405
KEYWORDS    EST.
SOURCE      human
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 280)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E.,
            Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel. 314 286 1800
            Fax. 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Insert Length: 728   Std Error: 0.00
            Seq primer: T7
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1..280
                     /organism="Homo sapiens"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Pac I - oligo(dT) primer
                     [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector.  Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
                     /clone="246872"
                     /clone_lib="Soares fetal liver spleen 1NFLS"
                     /sex="male"
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
     mRNA            1..280
BASE COUNT       80 a     45 c     75 g     61 t     19 others
ORIGIN

  Query Match             2.1%;  Score 26;  DB 25;  Length 280;
  Best Local Similarity   54.0%;  Pred. No. 3.29e-09;
  Matches   34;  Conservative    0;  Mismatches   26;  Indels    0;  Gaps    0;

Db         26 tcnggtancattanggagtcctcanctgtgnaagctanaagggatttttttaagcanaca 85
              || || || || || ||   ||        |||| | | || ||||  ||||||||  | |
Qy        310 TGGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATGTAAGGCCTTTTTTAAAAGAAGC 369

Db         86 gttnaaaataattacaaggtanantgnatataaaaccagnactgccggattgncaaaacn 145
               | |||      ||||      || |  ||   |    | ||    |||| |  ||||
Qy        370 ATTTAAGGATATAATGATTATATTTGTTCAGTTACAAATCAGTGTACAATCCATAAAAAC 429

Db        146 nagagaaagngttgttctttactgtcgttttcaaaatgtctaagtattggaatg 199
              | | | ||| | || |      ||| ||  |||  || |||     || |||| |
Qy        430 GGGTGTAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTACGAAGTGGGAATG 483


RESULT     5
LOCUS       DM36D7S        156 bp    DNA             STS       16-APR-1996
DEFINITION  D. melanogaster STS determined from European Mapping Project
            cosmid.
ACCESSION   Z70880
NID         g1263754
KEYWORDS    sequence tagged site.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Arthropoda;
            Tracheata; Insecta; Pterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 156)
  AUTHORS   European Drosophila Mapping,Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-APR-1996) Michael Ashburner, Department of Genetics,
            Downing St., Cambridge CB2 3EH, England
COMMENT     STS_name = Dm36D7S
            clone_name = 36D7
            STS_from_promoter = SP6
            vector_class = cosmid, Lorist 6
            origin_of_clone = Oregon-R
            in_situ_site_primary - 98C
            BLAST_program = BLASTN
            database_searched - EMBL
            database_version = 45.0 and updates till date_of_search
            date_of_search - 08-01-1996
            BLAST_program = BLASTX
            database_searched = SWISSPROT
            database_version - 32.0
            date_of_search = 15 12-1995.
FEATURES             Location/Qualifiers
     source          1..156
                     /organism="Drosophila melanogaster"
                     /strain="Oregon-R"
                     /clone="36D7"
BASE COUNT       28 a     37 c     30 g     27 t     34 others
ORIGIN

  Query Match             1.9%;  Score 24;  DB 36;  Length 156;
  Best Local Similarity  19.6%;  Pred. No. 1.72e-06;
  Matches    9;  Conservative   26;  Mismatches  11;  Indels   0;  Gaps   0;

Db      21 qsksrrqrrqrkrrqaqskssksrrqkkrkskrqssqqaqtrr 66
           |: ::::| :||:::: |:::::::: : | :: : :: ||| ::
Qy     597 GCCCCGAGTGCGGGAGCTGCTGCTGCACCCCTGAGCCCCGAGCAG 642


RESULT    6
ID    MM94013     standard; RNA; EST; 536 BP
AC    W20940;
NI    g1297857
DT    08-MAY-1996 (Rel. 47, Created)
DT    05-MAR-1997 (Rel. 51, Last updated, Version 2)
DE    mb96d08.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone 337263 5'
DE    similar to gb:M64497 APOLIPOPROTEIN AI REGULATORY PROTEIN-1
DE    (HUMAN); gb:U07635 Mus musculus ovalbumin upstream promoter
DE    transcription factor (MOUSE);.
KW    EST.
OS    Mus musculus (house mouse)
OC    Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
OC    Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
OC    Mus.
RN    [1]
RP    1-536
RA    Marra M., Hillier L., Allen M., Bowles M., Dietrich N.,
RA    Dubuque T., Geisel S., Kucaba T., Lacy M., Le M., Martin J.,
RA    Morris M., Schellenberg K., Steptoe M., Tan F., Underwood K.,
RA    Moore B., Theising B., Wylie T., Lennon G., Soares B., Wilson R.,
RA    Waterston R.,
RT    "The WashU-HHMI Mouse EST Project";
RL    Unpublished.
CC    Contact: Marra M/Mouse EST Project WashU-HHMI Mouse EST Project
CC    Washington University School of MedicineP 4444 Forest Park Parkway,
CC    Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810
CC    Email: mouseest@watson.wustl.edu This clone is available
CC    royalty free through LLNL ; contact the IMAGE Consortium
CC    (info@image.llnl.gov) for further information. MGI:218663 Seq
CC    primer: ETPrimer High quality sequence stop: 385.
FH    Key             Location/Qualifiers
FH
FT    source          1..536
FT                    /organism="Mus musculus"
FT                    /note "Vector: pT7T3D (Pharmacia) with a modified
FT                    polylinker, Site_1: Not I, Site_2: Eco RI, 1st strand cDNA
FT                    was primed with a Not I - oligo(dT) primer [5'
FT                    TGTTACCAATCTGAAGTGGGAGCGGCCGCATTTTTTTTTTTTTTTTTTT 3'],
FT                    double-stranded cDNA was size selected, ligated to Eco RI
FT                    adapters (Pharmacia), digested with Not I and cloned into
FT                    the Not I and Eco RI sites of a modified pT7T3 vector
FT                    (Pharmacia). Library went through one round of
FT                    normalization to a Cot = 5. Library constructed by Bento
FT                    Soares and M.Fatima Bonaldo. RNA was kindly provided by Dr

FT                    Minoru Ko (Wayne State University)."
```

```
FT                    /clone="337263"
FT                    /clone_lib="Soares mouse p3NMF19.5"
FT                    /dev_stage="19.5 dpc total fetus"
FT                    /lab_host="DH10B (ampicillin resistant)"
FT    mRNA            <1..>536
SQ    Sequence 536 BP; 113 A; 175 C; 164 G; 84 T; 0 other;

  Query Match             1.9%;  Score 24;  DB 97;  Length 536;
  Best Local Similarity  66.7%;  Pred. No. 1.72e-06;
  Matches   48;  Conservative    0;  Mismatches  24;  Indels   0;  Gaps   0;

Db     205 [illegible] 264
           ||| | |||||  |  |||  |||||     ||||||  |||||| ||     |||||    |
Qy     412 TGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTAC 471

Db     265 aaagtgggcatg 276
            ||||||| |||
Qy     472 GAAGTGGGAATG 483


RESULT    7
ID    HS1164487  standard; RNA; EST; 326 BP.
AC    AA249798;
NI    g1881068
DT    13-MAR-1997 (Rel. 51, Created)
DT    13-MAR-1997 (Rel. 51, Last updated, Version 1)
DE    111371.seq.F Human fetal heart, Lambda ZAP Express Homo sapiens
DE    cDNA 5'.
KW    EST.
OS    Homo sapiens (human)
OC    Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
OC    Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN    [1]
RP    1-326
RA    Liew C.C.;
RT    "cDNAs from human fetal heart (1997)";
RL    Unpublished.
CC    Contact: Liew CC Molecular Cardiology University of Toronto Banting
CC    Institute, 100 College St., Toronto, Ontario, M5G1L5 Tel:
CC    4169788758 Fax: 4169785650 Email: liewcc@utcc.utoronto.ca PCR
CC    PRimers FORWARD: 5' GCCAAGCTCGAAATTAACCCTCACTAAAGGG 3' BACKWARD: 5'
CC    CCAGTGAATTGTAATACGACTCACTATAGGGCG 3' Seq primer: 5'
CC    GAAATTAACCCTCACTAAAGGG 3'.
FH    Key             Location/Qualifiers
FH
FT    source          1..326
FT                    /organism="Homo sapiens"
FT                    /note="Vector: Lambda ZAP Express, Site_1: EcoRI, Site_2:
FT                    XhoI; mRNA was purified from human fetal hearts (8-10
FT                    weeks). cDNA was synthesized using a XhoI-Oligo dT
FT                    adaptor-primer. EcoRI adaptors were ligated, followed by
FT                    digestion with XhoI, for directional cloning into
FT                    predigested lambda ZAP Express."
FT                    /clone_lib="Human fetal heart, Lambda ZAP Express"
FT                    /lab_host="E. coli XL1-Blue"
FT    mRNA            <1..>326
SQ    Sequence 326 BP; 63 A; 87 C; 86 G; 90 T; 0 other;

  Query Match             1.8%;  Score 22;  DB 85;  Length 326;
  Best Local Similarity  78.9%;  Pred. No. 5.94e-04;
  Matches   30;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db     257 [illegible] 294
           |||| |||  ||  | |||  ||||||||||||||||
Qp     550 CGTCGGCACTTCTCTGTCTCCGCACAAGGCGGTACCCA 513


RESULT    8
LOCUS       AA249798       326 bp    mRNA            EST       11-MAR-1997
DEFINITION  111371.seq.F Human fetal heart, Lambda ZAP Express Homo sapiens
            cDNA 5'.
ACCESSION   AA249798
NID         g1881068
```

```
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 326)
  AUTHORS   Liew,C.C.
  TITLE     cDNAs from human fetal heart (1997)
  JOURNAL   Unpublished (1997)
COMMENT
            Contact: Liew CC
            Molecular Cardiology
            University of Toronto
            Banting Institute. 100 College St , Toronto, Ontario, M5G1L5
            Tel: 4169788758
            Fax: 4169785650
            Email: liewcc@utcc.utoronto.ca
            PCR PRimers
            FORWARD: 5' GCCAAGCTCGAAATTAACCCTCACTAAAGGG 3'
            BACKWARD: 5' CCAGTGAATTGTAATACGACTCACTATAGGGCG 3'
            Seq primer: 5' GAAATTAACCCTCACTAAAGGG 3'
FEATURES             Location/Qualifiers
     source          1..326
                     /organism="Homo sapiens"
                     /note="Vector: lambda ZAP Express; Site_1: EcoRI; Site_2:
                     XhoI; mRNA was purified from human fetal hearts (8-10
                     weeks).  cDNA was synthesized using a XhoI-Oligo dT
                     adaptor-primer.  EcoRI adaptors were ligated, followed by
                     digestion with XhoI, for directional cloning into
                     predigested lambda ZAP Express."
                     /clone_lib="Human fetal heart, Lambda ZAP Express"
                     /lab_host="E. coli XL1-Blue"
     mRNA            <1..>326
BASE COUNT       63 a     87 c     86 g     90 t
ORIGIN

  Query Match              1.8%;  Score 22;  DB 69;  Length 326;
  Best Local Similarity  78.9%;  Pred. No. 5.94e-04;
  Matches   30;  Conservative    0;  Mismatches   8;  Indels    0;  Gaps    0;

Db      257 cgtctgcagctcattatctggcacaaggcggtaccca 294
            ||:|| |||  :|  | |||  |||||||||||||||
Qy      550 CGTCGGCACTTCTCTGTCTCCGCACAAGGCGGTACCCA 513


RESULT     9
LOCUS       AA057786       382 bp    mRNA              EST       01-FEB-1997
DEFINITION  zf15e04.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone
            377022 3'.
ACCESSION   AA057786
NID         g1550437
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 382)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E.,
            Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1130   Std Error: 0.00
            Seq primer: -40M13 fwd. from Amersham
            High quality sequence stop: 272.
FEATURES             Location/Qualifiers
     source          1..382
                     /organism="Homo sapiens"
                     /note="Organ: heart; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCATCTTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library constructed by
                     M Fatima Bonaldo. This library was constructed from the
                     same fetus as the fetal lung library, Soares fetal lung
                     NbHL19W."
                     /clone="377022"
                     /clone_lib="Soares fetal heart NbHH19W"
                     /sex="unknown"
                     /dev_stage="19 weeks"
                     /lab_host="DH10B (ampicillin resistant)"
     mRNA            complement(<1..>382)
BASE COUNT       70 a    103 c    127 g     79 t      3 others
ORIGIN

  Query Match              1.8%;  Score 22;  DB 30;  Length 382;
  Best Local Similarity  78.9%;  Pred. No. 5.94e-04;
  Matches   30;  Conservative    0;  Mismatches   8;  Indels    0;  Gaps    0;

Db      145 gctgctggaggcacagggccccgggctgggagtgctca 182
            ||||||||:| |: | | ||||:| || |  ||:||:||
Qy      615 GCTGCTGGAGGCCCTGAGTCCCGAGTAGCTAGTGCTCA 652


RESULT    10
LOCUS       AA021604       399 bp    mRNA              EST       30-JAN-1997
DEFINITION  ze67c04.s1 Soares retina N2b4HR Homo sapiens cDNA clone 364038 3'
            similar to gb:X01677 GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE,
            LIVER (HUMAN);.
ACCESSION   AA021604
NID         g1485265
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 399)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E.,
            Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1473   Std Error: 0.00
            Seq primer: -40M13 fwd. from Amersham
            High quality sequence stop: 369.
FEATURES             Location/Qualifiers
     source          1..399
                     /organism="Homo sapiens"
                     /note="Organ: eye; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
```

```
                    strand cDNA was primed with a Not I - oligo(dT) primer [5'
                    TGTTACCAATCTGAAGTGGGAGCGGCCGCTTTTTTTTTTTTTTTTTT 3'],
                    double-stranded cDNA was size selected, ligated to Eco RI
                    adapters (Pharmacia), digested with Not I and cloned into
                    the Not I and Eco RI sites of a modified pT7T3 vector
                    (Pharmacia). The retinas were obtained from a 55 year old
                    Caucasian  and total cellular poly(A)+ RNA was extracted 6
                    hrs after their removal. The retina RNA  was kindly
                    provided by Roderick R. McInnes M.D. Ph.D. from the
                    University of Toronto. Library constructed by Bento
                    Soares and M.Fatima Bonaldo."
                    /clone="364038"
                    /clone_lib="Soares retina N2b4HR"
                    /sex="male"
                    /tissue_type="retina"
                    /dev_stage="55 year old"
                    /lab_host="DH10B (ampicillin resistant)"
     mRNA           complement(<1..>399)
BASE COUNT       74 a     92 c    132 g     98 t      3 others
ORIGIN

  Query Match             1.8%;  Score 22;  DB 21;  Length 399;
  Best Local Similarity  72.9%;  Pred. No. 5.94e-04;
  Matches   35;  Conservative    0;  Mismatches  13;  Indels    0;  Gaps    0;

Db     277 atgagcttgacaaagtggtccgttgagggcaatgccagccccagcgtt 324
           |||||||||| |    |||||  ||||| || |    |||||||| | ||
Qy     905 ATGAGCTTGCCGGGGTGGTCAATTGAGGGCCACATCAGCCCCATCATT 858


RESULT    11
ID    HSW0986     standard; RNA; EST; 424 BP.
AC    W96098;
NI    g1426005
DT    18-JUL-1996 (Rel. 48, Created)
DT    28-FEB-1997 (Rel. 51, Last updated, Version 2)
DE    ze08h09.r1 Soares fetal heart NbHH19W Homo sapiens cDNA clone
DE    358433 5' similar to gb:X52773 RETINOIC ACID RECEPTOR RXR-ALPHA
DE    (HUMAN);.
KW    EST.
OS    Homo sapiens (human)
OC    Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
OC    Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN    [1]
RP    1-424
RA    Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,
RA    Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,
RA    Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,
RA    Trevaskis E., Waterston R., Williamson A., Wohldmann P., Wilson R.;
RT    "The WashU Merck EST Project.",
RL    Unpublished.
CC    Contact: Wilson RK WashU-Merck EST Project Washington University
CC    School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis,
CC    MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:
CC    est@watson.wustl.edu This clone is available royalty-free through
CC    LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for
CC    further information. Insert Length: 1285 Std Error: 0.00 Seq
CC    primer: mob.REGA+ET.
FH    Key                 Location/Qualifiers
FH
FT    source              1..424
FT                        /organism="Homo sapiens"
FT                        /note="Organ: heart; Vector: pT7T3D (Pharmacia) with a
FT                        modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
FT                        strand cDNA was primed with a Not I - oligo(dT) primer [5'
FT                        TGTTACCAATCTGAAGTGGGAGCGGCCGCATCTTTTTTTTTTTTTTTTTT 3'],
FT                        double-stranded cDNAwas size selected, ligated to Eco RI
FT                        adapters (Pharmacia), digested with Not I and cloned into
FT                        the Not I and Eco RI sites of a modified pT7T3 vector
FT                        (Pharmacia). Library went through one round of
FT                        normalization to a Cot = 5. Library constructed by M.Fatim
  a
FT                        Bonaldo. This library was constructed from the same fetus
FT                        as the fetal lung library, Soares fetal lung NbHL19W."
FT                        /clone="358433"
FT                        /clone_lib="Soares fetal heart NbHH19W"
FT                        /sex="unknown"
FT                        /dev_stage="19 weeks"
FT                        /lab_host="DH10B (ampicillin resistant)"
FT    mRNA                <1..>424
SQ    Sequence 424 BP; 129 A; 92 C; 117 G; 85 T; 2 other;

  Query Match             1.8%;  Score 22;  DB 83;  Length 424;
  Best Local Similarity  62.8%;  Pred. No. 5.94e-04;
  Matches   54;  Conservative    0;  Mismatches  32;  Indels    0;  Gaps    0;

Db      36 cacatctgtgccatctgtggagacagatcctcaggaaagcactacggggtatacagttgt 95
           ||| |||| ||  ||||  | ||       | || ||| | ||||| || || |       |||
Qy     280 CACTTCTGCCCTGTCTGCAGGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGT 339

Db      96 gaaggctgcaaagggttcttcaagag 121
           ||||| || || |  || || || ||
Qy     340 GAAGGATGTAAGGCCTTTTTTAAAAG 365


RESULT    12
LOCUS       W96098        424 bp    mRNA            EST       29-NOV-1996
DEFINITION  ze08h09.r1 Soares fetal heart NbHH19W Homo sapiens cDNA clone
            358433 5' similar to gb:X52773 RETINOIC ACID RECEPTOR RXR-ALPHA
            (HUMAN);.
ACCESSION   W96098
NID         g1426005
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 424)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1285   Std Error: 0.00
            Seq primer: mob.REGA+ET.
FEATURES             Location/Qualifiers
     source          1..424
                     /organism="Homo sapiens"
                     /note="Organ: heart; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCATCTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library constructed by
                     M.Fatima Bonaldo. This library was constructed from the
                     same fetus as the fetal lung library, Soares fetal lung
                     NbHL19W."
                     /clone="358433"
                     /clone_lib="Soares fetal heart NbHH19W"
                     /sex="unknown"
                     /dev_stage="19 weeks"
                     /lab_host="DH10B (ampicillin resistant)"
     mRNA            <1..>424
BASE COUNT      128 a     92 c    117 g     85 t      2 others
ORIGIN

  Query Match             1.8%;  Score 22;  DB 2;  Length 424;
  Best Local Similarity  62.8%;  Pred. No. 5.94e-04;
  Matches   54;  Conservative    0;  Mismatches   32;  Indels    0;  Gaps    0;

Db        36 cacatctgtgccatctgtggagacagatcctcaggaaagcactacggggtatacagttgt 95
             ||| |||| ||  ||||  | ||       | || ||| | ||||| || || |      |||
Qy       280 CATTTCTGCGCTGTCTGCAGC[illegible] 339

Db        96 gaaggctgcaaagggttcttcaagag 121
             ||||| || || |   || || || ||
Qy       340 GAAGGATGTAAGGCCTTTTTTAAAAG 365


RESULT   13
LOCUS       AA024507      452 bp    mRNA            EST       30-JAN-1997
DEFINITION  ze76d08.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone
            364911 3'.
ACCESSION   AA024507
NID         g1489467
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 452)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E.,
            Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 783   Std Error: 0.00
            Seq primer: -40M13 fwd. from Amersham
            High quality sequence stop: 398.
FEATURES             Location/Qualifiers
     source          1..452
                     /organism="Homo sapiens"
                     /note="Organ: heart; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCATCTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library constructed by
                     M.Fatima Bonaldo. This library was constructed from the
                     same fetus as the fetal lung library, Soares fetal lung
                     NbHL19W."
                     /clone="364911"
                     /clone_lib="Soares fetal heart NbHH19W"
                     /sex="unknown"
                     /dev_stage="19 weeks"
                     /lab_host="DH10B (ampicillin resistant)"
     mRNA            complement(<1..>452)
BASE COUNT      112 a    105 c    132 g    101 t      2 others
ORIGIN

  Query Match             1.8%;  Score 22;  DB 28;  Length 452;
  Best Local Similarity  73.9%;  Pred. No. 5.94e-04;
  Matches   34;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Db       375 ccacggggccctcgtgtggggacgggtgcctgaggccctgagcccc 420
             ||||| | |||  ||| |||   | | |||  || ||||||||||||
Qy       591 CCACGCGCCCGAGTGGGGAGCTGCTGCTGGACGCCCTGAGCCCC 636


RESULT   14
LOCUS       AA196298      480 bp    mRNA            EST       22-JAN-1997
DEFINITION  [illegible] s1 Stratagene muscle 937209 Homo sapiens cDNA clone 627954
            3'.
ACCESSION   AA196298
NID         g1791889
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 480)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E.,
            Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40M13 fwd. from Amersham
            High quality sequence stop: 204.
FEATURES             Location/Qualifiers
     source          1..480
                     /organism="Homo sapiens"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI; Cloned unidirectionally.  Primer: Oligo dT. Skeletal
                     muscle from patient with malignant hyperthermia.  Average
                     insert size: 1.0 kb, Uni-ZAP XR Vector; -5' adaptor
                     sequence: 5' GAATTCGGCACGAG 3' -3' adaptor sequence: 5'
                     CTCGAGTTTTTTTTTTTTTTTTTT 3'"
                     /clone="627954"
                     /clone_lib="Stratagene muscle 937209"
                     /dev_stage="adult"
                     /lab_host="SOLR (kanamycin resistant)"
     mRNA            complement(<1..>480)
BASE COUNT       84 a    151 c    133 g    107 t      5 others
ORIGIN

  Query Match             1.9%;  Score 23;  DB 16;  Length 480;
  Best Local Similarity  69.0%;  Pred. No. 3.38e-05;
  Matches   40;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Db       203 [illegible] 260
Qy       250 [illegible] 193


RESULT   15
LOCUS       AA179181      498 bp    mRNA            EST       31-DEC-1996
DEFINITION  zp40b[illegible].r1 Stratagene HeLa cell s3 937216 Homo sapiens cDNA clone
            612473 5' similar to TR:E225526 E225526 ORF PZF1050. ;.
ACCESSION   AA179181
NID         g1760550
KEYWORDS    EST.
SOURCE      human.
```

```
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 498)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E.,
            Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 302.
FEATURES             Location/Qualifiers
     source          1..498
                     /organism="Homo sapiens"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI; Cloned unidirectionally.  Primer: Oligo dT. HeLa S3
                     epithelioid carcinoma cells grown to semi-confluency
                     without induction. Average insert size: 1.5 kb; Uni-ZAP XR
                     Vector  -5' adaptor sequence: 5' GAATTCGGCACGAG 3' -3'
                     adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTTTT 3'"
                     /clone="612473"
                     /clone_lib="Stratagene HeLa cell s3 937216"
                     /sex="female"
                     /dev_stage="HeLa S3 cell line"
                     /lab_host="SOLR (kanamycin resistant)"
     mRNA            <1..>498
BASE COUNT      113 a    164 c    129 g     86 t      6 others
ORIGIN

  Query Match              1.8%;  Score 22;  DB 11;  Length 498;
  Best Local Similarity  80.6%;  Pred. No. 5.94e-04;
  Matches    29;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Db      33 atccctggccagaacaagccaaggagccaagacgag 68
           ||||||| ||||||||||    ||||| |||| |||
Cp     936 ATCCCTGTCCAGAACAAGATCTGGAGCAAAGATGAG 901


Search completed: Thu Dec 18 19:15:04 1997
Job time : 365 secs.
```

```
***************************************************************************
                          MPSRCH                               (TM)
***************************************************************************

          Release 2 1D John F  Collins  Biocomputing Research Unit
          Copyright (c) 1993  1994, 1995  University of Edinburgh  U K
                 Distribution rights by IntelliGenetics, Inc

MPsrch_nn   n a  -  n a  database search, using Smith-Waterman algorithm

Run on:          Thu Dec 18 20 07 34 1997,  MasPar time 63 41 Seconds
                                            857 146 Million cell updates/sec
Tabular output not generated.

Title:           >US-08-826-361A-20
Description:     (1-1257) from US08826361A seq
Perfect Score:   1257
N.A. Sequence.       1 ATGAATTACAGCATTCCCAG.      .. GGCATGCGAGGTCTGCCTGA 1257
    Comp.              TACTTAATGTCGTAAGGGTC    ...   CCGTACGCTCCAGACGGACT

Scoring table:   TABLE default
                 Gap 6

Nmatch   STD :   Dbase 0;  Query 0

Searched:        81563 seqs, 21620618 bases x 2

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database:        n-issued
                   1:back1 2:51 3:52 4:53 5:54 6:55 7.56 8:PCT90 9.PCT91
                   10:PCT92 11:PCT93 12:PCT94 13:PCT95 14:PCT96

Statistics:      Mean 8.791;  Variance 4.451;  scale 1.975

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed
      and is derived by analysis of the total score distribution.

                                SUMMARIES
                 %
Result           Query
   No.   Score   Match Length DB  ID              Description             Pred. No.
-------------------------------------------------------------------------------
      1    135    10.7   4963  5  US-08-076-     Sequence 16, Applicati  2.38e-85
      2    135    10.7   4963  7  US-08-260-     Sequence 9, Applicatio  2.38e-85
c     3     71     5.6   2460  6  US-08-312-     Sequence 1, Applicatio  5.39e-35
      4     63     5.0   6244  5  US-08-076-     Sequence 15, Applicati  5.63e-29
      5     63     5.0   6244  7  US-08-260-     Sequence 8, Applicatio  5.63e-29
c     6     62     4.9   7219  7  US-08-232-     Sequence 14, Applicati  3.13e-28
      7     55     4.4   2989  1  5223606-1      Patent No. 5223606      4.54e-23
      8     53     4.2   3014  7  US-08-629-     Sequence 1, Applicatio  1.29e-21
      9     51     4.1   1576  1  5260432-1      Patent No. 5260432.     3.61e-20
     10     47     3.7    387  5  US-08-087-     Sequence 1, Applicati   2.60e-17
     11     47     3.7    387  6  US-08-561-     Sequence 1, Applicatio  2.60e-17
     12     42     3.3   2929 10  PCT-US92-0     Sequence 3, Applicati.  8.22e-14
     13     42     3.3   2940  1  5171671-1      Patent No. 5171671      8.22e-14
     14     42     3.3   3036 10  PCT-US92-0     Sequence 1, Applicati   8.22e-14
     15     38     3.0   1662  9  PCT-US91-0     Sequence 5, Applicatio  4.42e-11
     16     35     2.8    215  6  US-08-238-     Sequence 5, Applicatio  4.41e-09
     17     33     2.6   1952  6  US-08-333-     Sequence 1, Applicatio  8.90e-08
c    18     30     2.4    215  6  US-08-238-     Sequence 5, Applicatio  7.23e-06
     19     30     2.4   2095  9  PCT-US91-0     Sequence 3, Applicati   7.23e-06
     20     30     2.4   2129  5  US-07-952      Sequence 1, Applicatio  7.23e-06
     21     29     2.3   1866  9  PCT-US91-0     Sequence 1, Applicatio  3.03e-05
     22     28     2.2   2204  5  US-07-952      Sequence 2, Applicati.  1.25e-04
     23     28     2.2   2215  9  PCT-US91-0     Sequence 7, Applicatio  1.25e-04
     24     28     2.2   7468  6  US-08-333-     Sequence 11, Applicati  1.25e-04
     25     28     2.2   7219  7  US-08-232-     Sequence 14, Applicati  1.25e-04
     26     26     2.1   1649 12  PCT-US94-0     Sequence 1, Applicatio  1.99e-03
     27     26     2.1   1659  6  US-08-333-     Sequence 7, Applicatio  1.99e-03
     28     26     2.1   2304  7  US-08-464-     Sequence 1, Applicatio  1.99e-03
c    29     25     2.0     74 13  PCT-US95-1     Sequence 94, Applicati  7.66e-03
     30     25     2.0     75 13  PCT-US95-1     Sequence 99, Applicati  7.66e-03
c    31     25     2.0     81 13  PCT-US95-1     Sequence 92, Applicati  7.66e-03
     32     25     2.0     82 13  PCT-US95-1     Sequence 97, Applicati  7.66e-03
c    33     25     2.0     82 13  PCT-US95-1     Sequence 97, Applicati  7.66e-03
c    34     24     1.9     69  7  US-08-471-     Sequence 142, Applicat  2.89e-02
     35     24     1.9     74 13  PCT-US95-1     Sequence 94, Applicati  2.89e-02
c    36     24     1.9     74 13  PCT-US95-1     Sequence 100, Applicat  2.89e-02
c    37     24     1.9     75 13  PCT-US95-1     Sequence 99, Applicati  2.89e-02
     38     24     1.9     81 13  PCT-US95-1     Sequence 92, Applicati  2.89e-02
c    39     24     1.9     81 13  PCT-US95-1     Sequence 98, Applicati  2.89e-02
     40     24     1.9    105  5  US-07-865-     Sequence 13, Applicati  2.89e-02
     41     24     1.9   2009  6  US-08-333-     Sequence 9, Applicatio  2.89e-02
     42     23     1.8     74 13  PCT-US95-1     Sequence 100, Applicat  1.06e-01
c    43     23     1.8    105  5  US-07-865-     Sequence 13, Applicati  1.06e-01
     44     23     1.8   1577 13  PCT-US95-1     Sequence 14, Applicati  1.06e-01
     45     23     1.8   1787 13  PCT-US95-1     Sequence 6, Applicatio  1.06e-01

                                ALIGNMENTS

RESULT    1
ID   US-08-076-726-16 STANDARD; DNA; UNC; 4963 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 16, Application US/08076726.
CC   Sequence 16, Application US/08076726
CC   Patent No. 5464758
CC    GENERAL INFORMATION:
CC       APPLICANT:  Gossen, Manfred
CC       APPLICANT:  Bujard, Hermann
CC       TITLE OF INVENTION:  Tight Control of Gene Expression in
CC       TITLE OF INVENTION:  Eukaryotic Cells by Tetracycline responsive Promot
  ers
CC       NUMBER OF SEQUENCES:  16
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Sterne, Kessler, Goldstein & Fox
CC         STREET.  1100 New York Avenue, N.W.
CC         CITY.  Washington
CC         STATE:  District of Columbia
CC         COUNTRY:  United States of America
CC         ZIP:  20005-3934
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM.  PC-DOS/MS-DOS
CC         SOFTWARE   PatentIn Release #1 0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/076,726
CC         FILING DATE.  14-JUN-1993
CC         CLASSIFICATION.  435
CC       ATTORNEY,AGENT INFORMATION:
CC         NAME.  Esmond, Robert W.
CC         REGISTRATION NUMBER.  32,893
CC         REFERENCE/DOCKET NUMBER   1242.142.001
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE.  (202)371-2600
CC         TELEFAX.  (202)371-2540
CC         TELEX.  248636 SSK
CC     INFORMATION FOR SEQ ID NO:  16:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH.  4963 base pairs
CC         TYPE:  nucleic acid
CC         STRANDEDNESS.  both
CC         TOPOLOGY:  both
SQ   Sequence 4963 BP; 1259 A; 1300 C; 1278 G; 1126 T; 0 other;

  Query Match             10.7%,  Score 135,  DB 5;  Length 4963;
  Best Local Similarity  77.8%;  Pred. No  2.38e-85;
  Matches   189;  Conservative    0;  Mismatches   54;  Indels    0;  Gaps    0;

Db    1005 GGAGACTCGGCTACTGTGCAGTGTGCAATGACTATGCTTCAGGCTACCATTATGGAGTCTG 1064
           |||   |||  ||  |||  ||  ||  ||||    ||  ||  ||  ||  ||  ||  ||  |||||||||||
Qy     273 GGATGCTCACTTCTGGCGTGTCTGCAGCGATTACGGCATCGGGATATCACTATGGAGTCTG 332

Db    1065 GTCCTGTGAGGGCTGCAAGGCCTTCTTCAAGAGAAGTATTCAAGGACATAACGACTATAT 1124
           ||| |||||  ||  ||  ||||||||  ||  ||  |||||  ||||||||||||||  ||  |||||
Qy     333 GTCGTGTGAAGGATGTAAGGCCTTTTTAAAAGAAGCATTCAAGGACATAATGATTATAT 392

Db    1125 GTGTCCAGCCACCAACCAGTGCACCATTGATAAAAACAGGAGGAAGAGCTGCCAGGCCTG 1184
            ||||||||  ||  ||  |||||  ||  ||  |||||||||  ||  |  ||||||||||||||||||
Qy     393 TTGTCCAGCTACAAATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTG 452

Db    1185 CCGGCTCCGCAAATGCTACGAAGTGGGAATGATGAAAGGTGGGATACGAAAAGACCGAAC 1244
           |||  ||  ||  ||  ||  |||||||||||||||  ||||    ||||      ||  |  |||   ||  |
Qy     453 CCGACTTCGGAAGTGTTACGAAGTGGGAATGGTGAAGTGTGGCTCCCGGAGAGAGAGATG 512

Db    1245 AGG 1247
            ||
Qy     513 TGG 515


RESULT     2
ID   US-08-260-452-9 STANDARD; DNA; UNC; 4963 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 9, Application US/08260452.
CC   Sequence 9, Application US/08260452
CC   Patent No. 5650298
CC    GENERAL INFORMATION:
CC       APPLICANT:  Gossen, Manfred
CC       APPLICANT:  Bujard, Hermann
CC       APPLICANT.  Salfeld, Jochen
CC       APPLICANT:  Voss, Jeffrey
CC       TITLE OF INVENTION.  Tight Control of Gene Expression in Eucaryotic
CC       TITLE OF INVENTION:  Cells by Tetracycline-responsive Promoters
CC       NUMBER OF SEQUENCES:  10
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Lahive & Cockfield
CC         STREET:  60 State Street
CC         CITY:  Boston
CC         STATE:  Massachusetts
CC         COUNTRY.  USA
CC         ZIP.  02109 1875
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  ASCII text
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER.  US/08/260,452
CC         FILING DATE:
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER.  08/076,327
CC       ATTORNEY/AGENT INFORMATION
CC         NAME:  Giulio A. DeConti, Jr
CC         REGISTRATION NUMBER:  31,503
CC         REFERENCE/DOCKET NUMBER:  BBI-013CP
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (617) 227-7400
CC         TELEFAX:  (617) 227 5941
CC     INFORMATION FOR SEQ ID NO:  9:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  4963 base pairs
CC         TYPE.  nucleic acid
CC         STRANDEDNESS:  double
```

```
CC         TOPOLOGY:  circular
CC       MOLECULE TYPE:  DNA (genomic)
CC       ORIGINAL SOURCE:
CC         ORGANISM:  Human cytomegalovirus
CC       IMMEDIATE SOURCE:
CC         CLONE:  pUHD BGP4
SQ   Sequence 4963 BP; 1259 A; 1300 C; 1278 G; 1126 T; 0 other;

  Query Match             10.7%;  Score 135;  DB 7;  Length 4963;
  Best Local Similarity  77.8%,  Pred. No. 2.38e-95;
  Matches   189;  Conservative    0;  Mismatches   54;  Indels    0;  Gaps    0;

Db    1005 GGAGACTCGGCTACTGTGCAGTGTGCAATGACTATGCTTCAGGCTACCATTATGGAGTCTG 1064
           |||   |||  ||  |||  ||  ||  ||||    ||  ||  ||  ||  ||  ||  ||  |||||||||||
Qy     273 GGATGCTCACTTCTGGCGTGTCTGCAGCGATTACGGCATCGGGATATCACTATGGAGTCTG 332

Db    1065 GTCCTGTGAGGGCTGCAAGGCCTTCTTCAAGAGAAGTATTCAAGGACATAACGACTATAT 1124
           ||| |||||  ||  ||  ||||||||  ||  ||  |||||  ||||||||||||||  ||  |||||
Qy     333 GTCGTGTGAAGGATGTAAGGCCTTTTTAAAAGAAGCATTCAAGGACATAATGATTATAT 392

Db    1125 GTGTCCAGCCACCAACCAGTGCACCATTGATAAAAACAGGAGGAAGAGCTGCCAGGCCTG 1184
            ||||||||  ||  ||  |||||  ||  ||  |||||||||  ||  |  ||||||||||||||||||
Qy     393 TTGTCCAGCTACAAATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTG 452

Db    1185 CCGGCTCCGCAAATGCTACGAAGTGGGAATGATGAAAGGTGGGATACGAAAAGACCGAAG 1244
           |||  ||  ||  ||  ||  |||||||||||||||  ||||    ||||      ||  |  |||   ||  |
Qy     453 CCGACTTCGGAAGTGTTACGAAGTGGGAATGGTGAAGTGTGGCTCCCGGAGAGAGAGATG 512

Db    1245 AGG 1247
            ||
Qy     513 TGG 515


RESULT     3
ID   US-08-312-312A-1 STANDARD; DNA; UNC; 3460 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 1, Application US/08312312A.
CC   Sequence 1, Application US/08312312A
CC   Patent No. 5550024
CC    GENERAL INFORMATION:
CC       APPLICANT:  Rothschild, Max F.
CC       APPLICANT:  Tuggle, Christopher K.
CC       APPLICANT:  Jacobson, Carol D.
CC       APPLICANT:  Mileham, Alan J.
CC       APPLICANT:  Plastow, Graham S.
CC       APPLICANT:  Vaske, David A.
CC       TITLE OF INVENTION:  Genetic Markers for Pig Litter Size
CC       NUMBER OF SEQUENCES:  9
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Zarley, McKee, Thomte, Voorhees & Sease
CC         STREET:  801 Grand Suite 3200
CC         CITY:  Des Moines
CC         STATE:  Iowa
CC         COUNTRY:  United States
CC         ZIP:  50309
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE.  PatentIn Release #1.0, Version #1.30
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/312,312A
CC         FILING DATE:  26-SEP-1994
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER.  US 07/961,819
CC         FILING DATE.  19-OCT-1992
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 07/687,708
CC         FILING DATE:  19-APR-1991
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Nebel, Heidi S.
CC        REGISTRATION NUMBER:  37,719
CC        REFERENCE/DOCKET NUMBER:  BRLC-4
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  515-288-3667
CC        TELEFAX:  515-288-1338
CC    INFORMATION FOR SEQ ID NO:  1:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  3460 base pairs
CC        TYPE:  nucleic acid
CC        STRANDEDNESS:  double
CC        TOPOLOGY:  linear
CC      MOLECULE TYPE:  DNA (genomic)
CC      HYPOTHETICAL:  NO
CC      ANTI-SENSE:  NO
CC      FEATURE:
CC        NAME/KEY:  misc_feature
CC        LOCATION:  303..304
CC        OTHER INFORMATION:  /note= "Gap in DNA sequence of
CC        OTHER INFORMATION:  undetermined length"
CC      FEATURE:
CC        NAME/KEY:  misc_feature
CC        LOCATION:  1197..1198
CC        OTHER INFORMATION:  /note= "Gap in DNA sequence of
CC        OTHER INFORMATION:  undetermined length"
CC      FEATURE:
CC        NAME/KEY:  exon
CC        LOCATION:  534..600
SQ   Sequence 3460 BP; 1052 A; 765 C; 634 G; 1009 T; 0 other;

  Query Match             5.6%;  Score 71;  DB 6;  Length 3460;
  Best Local Similarity  83.3%;  Pred. No. 5.38e-35;
  Matches   95;  Conservative    0;  Mismatches  18;  Indels   1;  Gaps   1;

Db     490 TTCATCATGCCCACTTTGTAGCACTTGCCTAGCCGGCAGGG-TGATAGCTCTTTTCCTTG 548
           |||| ||| |||||||||||| ||||| || || ||||||||| || |||||||| | || |
Cp     488 TTCACCATTCCCACTTTGGTAACACTTCCGAAGTCGGTAGGCCTGGCAGCTCTTGCGCCGG 429

Db     549 TTCTTATCAATTGTGCACTGGTTGGTAGCTGGACACATGTAGTCATTATGTCCT 602
           || ||||| ||||| ||||| || |||||||||||| || || ||||||||||||
Cp     426 TTTTTATCGATTGTACACTGATTTGTAGCTGGACAAATATAATCATTATGTCCT 375


RESULT     4
ID   US-08-076-726-15 STANDARD; DNA; UNC; 6244 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 15, Application US/08076726
CC   Sequence 15, Application US/08076726
CC   Patent No. 5464758
CC    GENERAL INFORMATION:
CC      APPLICANT:  Gossen, Manfred
CC      APPLICANT:  Bujard, Hermann
CC      TITLE OF INVENTION:  Tight Control of Gene Expression in
CC      TITLE OF INVENTION:  Eucaryotic Cells by Tetracycline-responsive Promot
  ers
CC      NUMBER OF SEQUENCES:  16
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  Sterne, Kessler, Goldstein & Fox
CC        STREET:  1100 New York Avenue, N.W.
CC        CITY:  Washington
CC        STATE:  District of Columbia
CC        COUNTRY:  United States of America
CC        ZIP:  20005-3934
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  US/08/076,726
CC        FILING DATE:  14-JUN-1993
CC        CLASSIFICATION:  435
CC      ATTORNEY/AGENT INFORMATION:
```

```
CC        NAME:  Esmond, Robert W.
CC        REGISTRATION NUMBER:  32,893
CC        REFERENCE/DOCKET NUMBER:  0942.2490001
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  (202)371-2600
CC        TELEFAX:  (202)371-2540
CC        TELEX:  248636 SSK
CC    INFORMATION FOR SEQ ID NO:  15:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  6244 base pairs
CC        TYPE:  nucleic acid
CC        STRANDEDNESS:  both
CC        TOPOLOGY:  both
SQ   Sequence 6244 BP; 1483 A; 1749 C; 1610 G; 1402 T; 0 other;

  Query Match             5.0%;  Score 63;  DB 5;  Length 6244;
  Best Local Similarity  66.3%;  Pred. No. 5.63e-29;
  Matches  128;  Conservative    0;  Mismatches  65;  Indels   0;  Gaps   0;

Db    2365 TCTGTGGGGATGAAGCATCAGGCTGTCATTATGGTGTCCTCACCTGTGGGAGCTGTAAGG 2424
           ||||   | ||| | ||||| || | ||| ||||| |||     | ||||    | |||||||
Qy     293 TCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTCTGAAGGATGTAAGG 352

Db    2425 TTTTTTTTAAAAGGGGAATGGAAGGGCAGCATAACTATTTATCTGCTGGAAGAAATGACT 2484
            ||| ||||||||     ||    |||| ||   || | ||| | ||| | |   | |||| | |
Qy     353 CCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCACT 412

Db    2485 GCATTGTTGATAAAATCCCCAGGAAAAACTGCCCGGCCTGTCGGCTTAGAAAGTGCTGTC 2544
           | |    | ||||||| |||    | || |  ||||| ||| || || ||| | ||||| |
Qy     413 GTACAATCGATAAAAACGGGTGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTACG 472

Db    2545 AAGCTGGCATGGT 2557
           |||  || |||||
Qy     473 AAGTGGGAATGGT 485


RESULT     5
ID   US-08-260-452-8 STANDARD; DNA; UNC; 6244 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 8, Application US/08260452
CC   Sequence 8, Application US/08260452
CC   Patent No. 5650298
CC    GENERAL INFORMATION:
CC      APPLICANT:  Gossen, Manfred
CC      APPLICANT:  Bujard, Hermann
CC      APPLICANT:  Salfeld, Jochen
CC      APPLICANT:  Voss, Jeffrey
CC      TITLE OF INVENTION:  Tight Control of Gene Expression in Eucaryotic
CC      TITLE OF INVENTION:  Cells by Tetracycline-responsive Promoters
CC      NUMBER OF SEQUENCES:  10
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  Lahive & Cockfield
CC        STREET:  60 State Street
CC        CITY:  Boston
CC        STATE:  Massachusetts
CC        COUNTRY:  USA
CC        ZIP:  02109-1875
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  ASCII text
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  US/08/260,452
CC        FILING DATE:
CC        CLASSIFICATION:  435
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  08/076,327
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  Giulio A. DeConti, Jr.
CC        REGISTRATION NUMBER:  31,503
CC        REFERENCE/DOCKET NUMBER:  BBI-013CP
CC        TELECOMMUNICATION INFORMATION:
CC          TELEPHONE:  (617) 227 7400
CC          TELEFAX:  (617) 227-5941
CC      INFORMATION FOR SEQ ID NO:  8:
CC        SEQUENCE CHARACTERISTICS:
CC          LENGTH:  6244 base pairs
CC          TYPE:  nucleic acid
CC          STRANDEDNESS:  double
CC          TOPOLOGY:  circular
CC        MOLECULE TYPE:  DNA (genomic)
CC        ORIGINAL SOURCE:
CC          ORGANISM:  Human cytomegalovirus
CC          STRAIN:  Towne (hCMV)
CC        IMMEDIATE SOURCE:
CC          CLONE:  pUHD BGR3
SQ   Sequence 6244 BP, 1483 A, 1749 C, 1610 G, 1402 T, 0 other;

  Query Match              5.0%;  Score 63;  DB 7;  Length 6244;
  Best Local Similarity  66.3%;  Pred. No. 5.63e-29;
  Matches   129;  Conservative    0;  Mismatches  65;  Indels    0;  Gaps    0;

Db    2365 TCTGTGGCGATGAAGCATCAGCTTCTCATTATGGTGCCTCACCTCTGGGAGCTGTTAAGG 2424
           ||||   | ||| | ||||| || | ||| ||||| |||      | ||||     | |||||||
Qy     293 TCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATCTAAGG 352

Db    2425 TCTTCTTTAAAAGGCTAATGGAAGGGCAGCATAACTATTTATGTGCTGGAAGAAATGACT 2484
            ||| ||||||||     ||   |||| ||   || | ||| | ||| | |   | |||| | |
Qy     353 CCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCAGT 412

Db    2485 GCATTCTTGATAAAATCCGCAGGAAAAACTGCCCGGCCGTGTCCGTTAGAAAGTGCTGTC 2544
           | |    | ||||||| |||   | || | ||||| ||| || || ||| | ||||| |
Qy     413 GTACAATCGATAAAAACGGCGGCAAGAGCTGCCAGGTCTGCCGACTTCGGAAGTGTTACG 472

Db    2545 AAGCTGGCATGGT 2557
           |||  || |||||
Qy     473 AAGTGGGAATGGT 485


RESULT     6
ID   US-08-232-463-14 STANDARD; DNA; UNC; 7218 BP
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 14, Application US/08232463.
CC   Sequence 14, Application US/08232463
CC   Patent No. 5670367
CC    GENERAL INFORMATION:
CC       APPLICANT:  DORNER, F.
CC       APPLICANT:  SCHEIFLINGER, F.
CC       APPLICANT:  FALKNER, F. G.
CC       TITLE OF INVENTION:  RECOMBINANT FOWLPOX VIRUS
CC       NUMBER OF SEQUENCES:  52
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Foley & Lardner
CC         STREET:  1800 Diagonal Road, Suite 500
CC         CITY:  Alexandria
CC         STATE:  VA
CC         COUNTRY:  USA
CC         ZIP:  22313-0299
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/232,463
CC         FILING DATE:
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/935,313
CC         FILING DATE:
CC         APPLICATION NUMBER:  EP 91 114 300.6
CC         FILING DATE:  26-AUG-1991
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  BENT, Stephen A.
CC         REGISTRATION NUMBER:  29,768
CC         REFERENCE/DOCKET NUMBER:  30472/114 IMMU
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (703)836-9300
CC         TELEFAX:  (703)683-4109
CC         TELEX:  899149
CC     INFORMATION FOR SEQ ID NO:  14:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  7218 base pairs
CC         TYPE:  nucleic acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       IMMEDIATE SOURCE:
CC         CLONE:  pTZgpt-Fls
SQ   Sequence 7218 BP, 1944 A, 1491 C, 1490 G, 1425 T, 868 other;

  Query Match              4.9%;  Score 62;  DB 7;  Length 7218;
  Best Local Similarity   1.4%;  Pred. No. 3.13e-28;
  Matches     5;  Conservative  205;  Mismatches 148;  Indels    0;  Gaps    0;

Db    1083 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1142
           ::::: : :::       : : :   :: :: ::::::::  :::: ::   : :   : :   ::
Qy     612 CTCCCCACTCGGCCCCCCGTGCCCCCCACTTCTCTTCCCCTTCCCGGCACAGTGCAGCTG 553

Db    1143 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1202
           ::: ::   : ::::::: ::::: : :       :   : ::: : :::::::::::      ::
Qy     552 CTCCTCGGCACTTCTCTGTCTCCGCACAAGGCGGTACCCACATCTCTCTCTCCGGGAGCC 493

Db    1203 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1262
             : :... .. ..... ...: .   . ....:       ::   :     :::   :   ::::: :
Qy     492 ACACTTCACCATTCCCACTTCGTAACACTTCCGAAGTCGGCAGCCCTGGCAGCTCTTGCG 433

Db    1263 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1322
           ::   ::::: ::   :: : : ::   ::: :   ::      :    : :   :: :: : :::::
Qy     432 CCGGTTTTTATCGATTGTACACTGATTTGTAGCTGGACAAATATAATCATTATGTCCTTG 373

Db    1323 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1382
              : :::::::::            :::: : :::::: : :   ::    :::: :   :   : :::
Qy     372 AATGCTTCTTTTAAAAAAGGCCTTACATCCTCCACACCACCACACTCCATAGTGATATCC 313

Db    1383 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYGTACCA 1440
            .  .  .  .  .. .. .     :  .  .  .     .    . .........    ...| ||||
Qy     312 CGATCCTTAATCCTCGTAGACACTCCAGAAGTGAGCATCCCTCTTTGAACCTGGACCA 255


RESULT     7
ID   5223606-1 STANDARD; DNA; UNC; 3238 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Patent No. 5223606.
CC  Patent No. 5223606
CC       APPLICANT:  BLAUDIN DE THE, HUGHES;MARCHIO, AGNES;TIOLLAIS,
CC  PIERRE;DEJEAN, ANNE
CC       TITLE OF INVENTION: STEROID/THYROID HORMONE RECEPTOR-RELATED
CC  PROTEIN INAPPROPRIATELY EXPRESSED IN HUMAN HEPATOCELLULAR CARCINOMA
CC       NUMBER OF SEQUENCES: 11
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/134 140
CC         FILING DATE: 17-DEC-1987
CC       PRIOR APPLICATION DATA:
CC  SEQ ID NO:1:
CC         LENGTH: 2989
SQ   Sequence 3238 BP; 900 A; 648 C; 600 G; 841 T; 249 other;

  Query Match              4.4%;  Score 55;  DB 1;  Length 2989;
  Best Local Similarity  65.0%;  Pred. No. 4.54e-23;
  Matches   130;  Conservative    0;  Mismatches  69;  Indels    1;  Gaps    1;

Db     560 CTGCTTCGTCTGCCAGGACAAATCATCAGGGTACCACTATGGGGTCAGCGCCTGTGAGGG 619
           ||||    ||||||    ||   |    |||| || || |||||||| ||| |    | ||||| ||
Qy     285 CTGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGG 344
```

```
Db     620 ATGTAAGGGCTTTTTCCGCAGAAGTATTCAGAAGAATATGATTTACACTTGTCACCGAGA 679
           |||||||| ||||||    ||||| |||||     |||    ||| | |||||
Qy     345 ATGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTAC 404

Db     680 TAAGAACTGTGTTATTAATAAAAGTCACCAGGAATCGATGCCAATACTGTCGACTCCAGA 739
            ||  | |||   ||  ||||||  |  | | ||  | |||||   |||  ||||| | ||
Qy     405 AAATCAGTGTACAATCGATAAAAACCGGC-GCAAGAGCTGCCAGGCCTGCCGACTTCGGA 463

Db     740 AGTGCTTTGAAGTGGGAATG 759
           |||| |  ||||||||||||
Qy     464 AGTGTTACGAAGTGGGAATG 483


RESULT    8
ID   US-08-629-939-1 STANDARD; DNA; UNC; 3014 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 1, Application US/08629939.
CC   Sequence 1, Application US/08629939
CC   Patent No. 5645995
CC     GENERAL INFORMATION:
CC       APPLICANT:  Kieback, Dirk G.
CC       TITLE OF INVENTION:  METHODS FOR DIAGNOSING AN INCREASED
CC       TITLE OF INVENTION:  RISK OF BREAST OR OVARIAN CANCER
CC       NUMBER OF SEQUENCES:  14
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  SUGHRUE, MION, ZINN, MACPEAK & SEAS
CC         STREET:  2100 Pennsylvania Avenue, N.W., Suite 800
CC         CITY:  Washington, D.C
CC         STATE:  D.C.
CC         COUNTRY:  U S A
CC         ZIP:  20037
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/629,939
CC         FILING DATE:  12-APRIL-1996
CC         CLASSIFICATION:  435
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  KIT, Gordon
CC         REGISTRATION NUMBER:  30,764
CC         REFERENCE/DOCKET NUMBER:  A-6612
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (202) 293-7060
CC         TELEFAX:  (202) 293-7860
CC     INFORMATION FOR SEQ ID NO:  1:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  3014 base pairs
CC         TYPE:  nucleic acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  cDNA
CC       HYPOTHETICAL:  NO
SQ   Sequence 3014 BP; 623 A; 939 C; 827 G; 625 T; 0 other;

  Query Match              4.2%;  Score 53;  DB 7;  Length 3014;
  Best Local Similarity  63.7%;  Pred. No. 1.29e-21;
  Matches  123;  Conservative    0;  Mismatches  70;  Indels    0;  Gaps    0;

Db    1881 TCTGTGGGGATGAAGCATCAGGCTGTCATTATGGTGTCCTTACCTGTGGGAGCTGTAAGG 1940
           ||||  | ||| | | ||| || | ||| ||||| |||    | ||||   | |||||||
Qy     293 TCTGCAGTGATTACGTATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATGTAAGG 352

Db    1941 TCTTTTTTAAGAGGGTAATGGAAGGGCAGTACAACTACTTATGTCCTGGAAGAAATGACT 2000
            ||| ||||| ||     ||  |||| ||  |   |  | ||  | ||| | |   | ||||
Qy     353 CCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCAGT 412

Db    2001 GTATCGTCGATAAAATCCTTAGAAAAAACTGCCCAGCATGTTGCCTTAGAAAGTGCTGTC 2060
            | |   | |||||||| |||    | || | |||||  || || || ||| | ||||| |
Qy     413 GTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTACG 472
```

```
Db    2061 AGGCTGGCATGGT 2073
           | |  || |||||
Qy     473 AAGTGGGAATGGT 485


RESULT    9
ID   5260432-1 STANDARD; DNA; UNC; 1707 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Patent No. 5260432.
CC   Patent No. 5260432
CC       APPLICANT: TAKAKU, FUMIMARO;ISHIKAWA, TAKASHI;IMAWARI, MICHIO;
CC   EVANS, RONALD M.;UMESONO, KAZUHIKO
CC       TITLE OF INVENTION: HUMAN GAMMA RETINOIC ACID RECEPTOR DNA
CC       NUMBER OF SEQUENCES: 4
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/370,407
CC         FILING DATE: 22-JUN-1989
CC   SEQ ID NO.1.
CC         LENGTH: 1576
SQ   Sequence 1707 BP; 346 A; 478 C; 467 G; 285 T; 131 other;

  Query Match              4.1%;  Score 51;  DB 1;  Length 1576;
  Best Local Similarity  65.8%;  Pred. No. 3.61e-20;
  Matches  127;  Conservative    0;  Mismatches  64;  Indels    2;  Gaps    2;

Db     473 GTGTGCAATGACAAGTCCTCTGGCTACCACTATGGGGTCAGCCTTTCTGAAGCGCTGCAAG 532
           || ||||  ||  |    | || || || |||||||| |||  | || |||||||| || |||
Qy     292 GTCTGCAGTGATTACGTATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATGTAAG 351

Db     533 GGCTTCTTTCGCCGAAGCATCCAGAAGAACATGGTGTACACGTGTCACCGCGACAAAA-A 591
           | ||| |||     ||||||| ||        | |  |  || |  |||| | || |||||  |
Qy     352 GCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTC-CAGCTACAAATCA 410

Db     592 CTGTATCATCAACAAGGTGACTAGGAATCGCTGCCAGTACTGCCGCTTACAGAAGTGCTT 651
             ||||  ||| | ||        | ||  ||||||||  |||||| || | |||||| |
Qy     411 GTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTA 470

Db     652 CGAAGTGGGCATG 664
           ||||||||| |||
Qy     471 CGAAGTGGGAATG 483


RESULT   10
ID   US-08-087-151-1 STANDARD; DNA; UNC; 387 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 1, Application US/08087151.
CC   Sequence 1, Application US/08087151
CC   Patent No. 5468624
CC     GENERAL INFORMATION:
CC       APPLICANT:  E. Brad Thompson
CC       APPLICANT:  Lynne V. Nazareth
CC       TITLE OF INVENTION:  ANTI-TUMOR ACTIVITY OF A MODIFIED FRAGMENT
CC       TITLE OF INVENTION:  OF THE GLUCOCORTICOID RECEPTOR
CC       NUMBER OF SEQUENCES:  3
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Arnold, White & Durkee
CC         STREET:  P.O. Box 4433
CC         CITY:  Houston
CC         STATE:  Texas
CC         COUNTRY:  USA
CC         ZIP:  77210
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  WordPerfect 5.1
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/087,151
CC         FILING DATE:  July 7, 1993
CC         CLASSIFICATION:  435
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Hodgins, Daniel S.
CC         REGISTRATION NUMBER:  31,026
CC         REFERENCE/DOCKET NUMBER:  UTSB 141/HOD
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  512-320-7200
CC         TELEFAX:  512-474-7577
CC         TELEX:  No. 5468624 Applicable
CC     INFORMATION FOR SEQ ID NO:  1:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  387 base pairs
CC         TYPE:  nucleic acid
CC         STRANDEDNESS:  double
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  genomic DNA
SQ   Sequence 387 BP; 109 A; 92 C; 88 G; 98 T; 0 other;

  Query Match              3.7%;  Score 47;  DB 5;  Length 387;
  Best Local Similarity  63.0%;  Pred. No. 2.60e-17;
  Matches   126;  Conservative    0;  Mismatches  73;  Indels    1;  Gaps    1;

Db      188 TCTGCCTGGTGTGCTCTGATGAAGCTTCAGGATGTCATTATGGAGTCTTAACTTGTGGAA 247
            |||||   || |||   ||| | || || |||| ||| ||||||||||   | |||| |
Qy      284 TCTGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAG 343

Db      248 GCTGTAAAGTTTTCTTCAAAAGAGCAGTGGAAGGACAGCACAATTACCTATGTGCTGGAA 307
            | ||||| |  || || ||||||    |  |||||||  |  ||||  | ||| | |  |
Qy      344 GATGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTA 403

Db      308 GGAATGATTGCATCATCGCGATAAAATT-CGAAAAAACTGCCCAGCATGCCGCTATCGAA 366
              ||| | || |  ||||  | |||    || || | |||||  || |||||   ||| |
Qy      404 CAAATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGA 463

Db      367 AATGTCTTCAGGCTGGAATG 386
            | |||    | |  ||||||
Qy      464 AGTGTTACGAAGTGGGAATG 483


RESULT   11
ID   US-08-561-072-1 STANDARD; DNA; UNC; 387 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 1, Application US/08561072
CC   Sequence 1, Application US/08561072
CC   Patent No. 5571791
CC    GENERAL INFORMATION:
CC       APPLICANT:  E. Brad Thompson
CC       APPLICANT:  Lynne V. Nazareth
CC       TITLE OF INVENTION:  ANTI-TUMOR ACTIVITY OF A MODIFIED FRAGMENT
CC       TITLE OF INVENTION:  OF THE GLUCOCORTICOID RECEPTOR
CC       NUMBER OF SEQUENCES:  3
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Arnold, White & Durkee
CC         STREET:  P.O. Box 4433
CC         CITY:  Houston
CC         STATE:  Texas
CC         COUNTRY:  USA
CC         ZIP:  77210
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  WordPerfect 5.1
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/561,072
CC         FILING DATE:  20-NOV-1995
CC         CLASSIFICATION:  530
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  08/087,151
CC         FILING DATE:  July 7, 1993
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Hodgins, Daniel S.
CC         REGISTRATION NUMBER:  31,026
```

```
CC         REFERENCE/DOCKET NUMBER:  UTSB:141/HOD
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  512-320-7200
CC         TELEFAX.  512-474-7577
CC         TELEX:  No. 5571791 Applicable
CC     INFORMATION FOR SEQ ID NO:  1:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  387 base pairs
CC         TYPE:  nucleic acid
CC         STRANDEDNESS:  double
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  genomic DNA
SQ   Sequence 387 BP; 109 A; 92 C; 88 G; 98 T; 0 other;

  Query Match              3.7%;  Score 47;  DB 6;  Length 387;
  Best Local Similarity  63.0%;  Pred. No. 2.60e-17;
  Matches   126;  Conservative    0;  Mismatches  73;  Indels    1;  Gaps    1;

Db      188 TCTGCCTGGTGTGCTCTGATGAAGCTTCAGGATGTCATTATGGAGTCTTAACTTGTGGAA 247
            |||||   || |||   ||| | || || |||| ||| ||||||||||   | |||| |
Qy      284 TCTGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAG 343

Db      248 GCTGTAAAGTTTTCTTCAAAAGAGCAGTGGAAGGACAGCACAATTACCTATGTGCTGGAA 307
            | ||||| |  || || ||||||    |  |||||||  |  ||||  | ||| | |  |
Qy      344 GATGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTA 403

Db      308 GGAATGATTGCATCATCGCGATAAAATT-CGAAAAAACTGCCCAGCATGCCGCTATCGAA 366
              ||| | || |  ||||  | |||    || || | |||||  || |||||   ||| |
Qy      404 CAAATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGA 463

Db      367 AATGTCTTCAGGCTGGAATG 386
            | |||    | |  ||||||
Qy      464 AGTGTTACGAAGTGGGAATG 483


RESULT   12
ID   PCT-US92-02320A-3 STANDARD; DNA; UNC; 2928 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 3, Application PC/US9202320A
CC   Sequence 3, Application PC/US9202320A
CC    GENERAL INFORMATION:
CC       APPLICANT:  Sloan-Kettering Institute, For Cancer Research
CC       TITLE OF INVENTION:  METHODS FOR DETECTION AND TREATMENT OF CANCER
CC       NUMBER OF SEQUENCES:  6
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  COOPER & DUNHAM
CC         STREET:  30 ROCKEFELLER PLAZA
CC         CITY:  NEW YORK
CC         STATE:  NEW YORK
CC         COUNTRY:  U.S.A.
CC         ZIP:  10112
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.24
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US92/02320A
CC         FILING DATE:
CC         CLASSIFICATION:
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 675,084
CC         FILING DATE:  22-MAR-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 673,838
CC         FILING DATE:  22-MAR-1991
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  WHITE, JOHN P
CC         REGISTRATION NUMBER:  28,678
CC         REFERENCE/DOCKET NUMBER:  38694-PCT
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 977-9550
CC        TELEFAX: (212) 644-0525
CC        TELEX: (212) 422523 COOP UI
CC    INFORMATION FOR SEQ ID NO: 3:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH: 2928 base pairs
CC        TYPE: nucleic acid
CC        STRANDEDNESS: single
CC        TOPOLOGY: linear
CC      MOLECULE TYPE: cDNA
CC      IMMEDIATE SOURCE:
CC        CLONE: hRAR ALPHA
CC      FEATURE:
CC        NAME/KEY: CDS
CC        LOCATION: 103..1488
CC        OTHER INFORMATION:
SQ   Sequence 2928 BP; 582 A; 1011 C; 789 G; 546 T; 0 other;

  Query Match             3.3%;  Score 42;  DB 10;  Length 2928;
  Best Local Similarity   60.6%;  Pred. No. 8.22e-14;
  Matches  120;  Conservative    0;  Mismatches   78;  Indels    0;  Gaps    0;

Db    364 TGCTTTGTCTGTCAGGACAAGTCCTCAGGCTACCACTATGGGGTCAGCGCCTGTGAGGGC 423
          |||  ||||||    ||  |  | || || || |||||||| ||| |  | ||||| ||
Qy    286 TGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGA 345

Db    424 TGCAAGGGCTTCTTCCGCCGCAGCATCCAGAAGAACATGGTGTACACGTGTCACCGGGAC 483
          || |||| ||| ||     | ||||| ||     | |  |  || |  ||||
Qy    346 TGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACA 405

Db    484 AAGAACTGCATCATCAACAAGGTGACCCGGAACCGCTGCCAGTACTGCCGACTGCAGAAG 543
          ||  | || |  ||| | ||       || ||  ||||||||  ||||||||| | ||||
Qy    406 AATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAG 465

Db    544 TGCTTTGAAGTGGGCATG 561
          || |  |||||||| |||
Qy    466 TGTTACGAAGTGGGAATG 483

RESULT   13
ID   5171671-1 STANDARD; DNA; UNC; 3185 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Patent No. 5171671.
CC  Patent No. 5171671
CC      APPLICANT: EVANS, RONALD M.;ONG, ESTELITA S.;SEGUI,
CC  PRUDIMAR S.;THOMPSON, CATHERINE C.;UEMSONO, KAZUHIKO
CC  GUGUERE, VINCENT
CC      TITLE OF INVENTION: RETINOIC ACID RECEPTOR COMPOSITION
CC      NUMBER OF SEQUENCES: 2
CC      CURRENT APPLICATION DATA
CC        APPLICATION NUMBER  US/07/546,256
CC        FILING DATE: 06-AUG-1990
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER: 276,536
CC        FILING DATE: 30-NOV-1988
CC        APPLICATION NUMBER: 128,331
CC        FILING DATE: 02-DEC-1987
CC  SEQ ID NO:1:
CC        LENGTH: 2940
SQ   Sequence 3185 BP; 594 A; 1010 C; 790 G; 546 T; 245 other;

  Query Match             3.3%;  Score 42;  DB 1;  Length 2940;
  Best Local Similarity   60.6%;  Pred. No. 8.22e-14;
  Matches  120;  Conservative    0;  Mismatches   78;  Indels    0;  Gaps    0;

Db    364 TGCTTTGTCTGTCAGGACAAGTCCTCAGGCTACCACTATGGGGTCAGCGCCTGTGAGGGC 423
          |||  ||||||    ||  |  | || || || |||||||| ||| |  | ||||| ||
Qy    286 TGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGA 345

Db    424 TGCAAGGGCTTCTTCCGCCGCAGCATCCAGAAGAACATGGTGTACACGTGTCACCGGGAC 483
          || |||| ||| ||     | ||||| ||     | |  |  || |  ||||
Qy    346 TGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACA 405

Db    484 AAGAACTGCATCATCAACAAGGTGACCCGGAACCGCTGCCAGTACTGCCGACTGCAGAAG 543
          ||  | || |  ||| | ||       || ||  ||||||||  ||||||||| | ||||
Qy    406 AATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAG 465

Db    544 TGCTTTGAAGTGGGCATG 561
          || |  |||||||| |||
Qy    466 TGTTACGAAGTGGGAATG 483

RESULT   14
ID   PCT-US92-02320A-1 STANDARD; DNA; UNC; 3036 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 1, Application PC/TUS9202320A.
CC   Sequence 1, Application PC/TUS9202320A
CC    GENERAL INFORMATION:
CC      APPLICANT: Sloan-Kettering Institute, For Cancer Research
CC      TITLE OF INVENTION: METHODS FOR DETECTION AND TREATMENT OF CANCER
CC      NUMBER OF SEQUENCES: 6
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE: COOPER & DUNHAM
CC        STREET: 30 ROCKEFELLER PLAZA
CC        CITY: NEW YORK
CC        STATE: NEW YORK
CC        COUNTRY: U.S.A
CC        ZIP: 10112
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE: Floppy disk
CC        COMPUTER: IBM PC compatible
CC        OPERATING SYSTEM: PC-DOS/MS-DOS
CC        SOFTWARE: PatentIn Release #1.24
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER: PCT/US92/02320A
CC        FILING DATE:
CC        CLASSIFICATION:
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER: US 675,084
CC        FILING DATE: 22-MAR-1991
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER: US 673,838
CC        FILING DATE: 22-MAR-1991
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME: WHITE, JOHN P
CC        REGISTRATION NUMBER: 28,678
CC        REFERENCE/DOCKET NUMBER: 38694-PCT
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE: (212) 977-9550
CC        TELEFAX: (212) 644-0525
CC        TELEX: (212) 422523 COOP UI
CC    INFORMATION FOR SEQ ID NO: 1:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH: 3036 base pairs
CC        TYPE: nucleic acid
CC        STRANDEDNESS: single
CC        TOPOLOGY: linear
CC      MOLECULE TYPE: cDNA
CC      IMMEDIATE SOURCE:
CC        CLONE: MYL-RAR
CC      FEATURE:
CC        NAME/KEY: CDS
CC        LOCATION: 67..2457
CC        OTHER INFORMATION:
SQ   Sequence 3036 BP; 605 A; 1018 C; 890 G; 523 T; 0 other;

  Query Match             3.3%;  Score 42;  DB 10;  Length 3036;
  Best Local Similarity   60.6%;  Pred. No. 8.22e-14;
  Matches  120;  Conservative    0;  Mismatches   78;  Indels    0;  Gaps    0;

Db   1333 TGCTTTGTCTGTCAGGACAAGTCCTCAGGCTACCACTATGGGGTCAGCGCCTGTGAGGGC 1392
          |||  ||||||    ||  |  | || || || |||||||| ||| |  | ||||| ||
Qy    286 TGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGA 345

Db   1393 TGCAAGGGCTTCTTCCGCCGCAGCATCCAGAAGAACATGGTGTACACGTGTCACCGGGAC 1452
```

```
           || |||| ||| ||     | ||||| ||     | |  |  || |  ||||
Qy     346 TGTAAGGCCTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACA 405

Db    1453 AAGAACTGCATCATCAACAAGGTGACCCGGAACCGCTGCCAGTACTGCCGACTGCAGAAG 1512
           ||  | || |  ||| | ||       || ||   ||||||||  ||||||||| | ||||
Qy     406 AATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAG 465

Db    1513 TGCTTTGAAGTGGGCATG 1530
           || |  |||||||| |||
Qy     466 TGTTACGAAGTGGGAATG 483

RESULT    15
ID    PCT-US91-00399-5 STANDARD; DNA; UNC; 1662 BP.
AC    xxxxxx
DT    01-JAN-1900
DE    Sequence 5, Application PC/TUS9100399.
CC    Sequence 5, Application PC/TUS9100399
CC     GENERAL INFORMATION:
CC       APPLICANT:  Mangelsdorf, Dr., David J.
CC       APPLICANT:  Evans Dr., Ronald M.
CC       TITLE OF INVENTION:  RETINOID RECEPTOR COMPOSITIONS AND
CC       TITLE OF INVENTION:  METHODS
CC       NUMBER OF SEQUENCES:  7
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  FITCH, EVEN, TABIN & FLANNERY
CC         STREET:  135 South LaSalle Street, Suite 900
CC         CITY:  Chicago
CC         STATE:  Illinois
CC         COUNTRY:  USA
CC         ZIP:  60603
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US91/00399
CC         FILING DATE:  19910122
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/478,071
CC         FILING DATE:  09-FEB-1990
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Reiter, Mr., Stephen E.
CC         REGISTRATION NUMBER:  31192
CC         REFERENCE/DOCKET NUMBER:  50852
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (619) 552-1311
CC         TELEFAX:  (619) 552-0095
CC         TELEX:  20 6566 PATLAW CGO
CC     INFORMATION FOR SEQ ID NO:  5:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  1662 base pairs
CC         TYPE:  NUCLEIC ACID
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  DNA (genomic)
CC       IMMEDIATE SOURCE:
CC         CLONE:  RXR MOUSE GAMMA
CC       FEATURE:
CC         NAME/KEY:  CDS
CC         LOCATION:  join(123..1514)
SQ    Sequence 1662 BP; 397 A; 483 C; 441 G; 341 T; 0 other;

  Query Match             3.0%;  Score 38;  DB 9;  Length 1662;
  Best Local Similarity   60.1%;  Pred. No. 4.42e-11;
  Matches  113;  Conservative    0;  Mismatches   75;  Indels    0;  Gaps    0;

Db     531 CACATCTGTGCCATCTGTGGGGACAGATCCTCAGGGAAGCACTACGGTGTGTACAGCTGT 590
           ||| |||| ||  ||||  | ||     | || ||  |  ||||| || || |     |||
Qy     280 CACTTCTGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGT 339
```

```
Db     591 GAAGGTTGCAAAGGCTTCTTCAAAAGGACCATCAGGAAAGATCTCATCTACACCTGTCGG 650
           ||||| || || | ||| || ||||| | |||       | ||      || |  ||||
Qy     340 GAAGGATGTAAGGCCTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCA 399

Db     651 GATAACAAAGATTGTCTCATCGACAAGCGCCAGCGCAACCGCTGCCAGTACTGTCGCTAC 710
           | ||  ||  |  |||   ||||| ||    || ||||||  ||||||||  ||| ||
Qy     400 GCTACAAATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTT 459

Db     711 CAGAAGTG 718
           | ||||||
Qy     460 CGGAAGTG 467

Search completed: Thu Dec 18 20:08:40 1997
Job time : 66 secs.
```

```
***********************************************************************

                                                      (TM)

***********************************************************************

         Release 2.1D John F. Collins, Biocomputing Research Unit.
         Copyright (c) 1993, 1994, 1995  University of Edinburgh, U.K.
                  Distribution rights by IntelliGenetics, Inc.

MPsrch_pp   protein - protein database search, using Smith-Waterman algorithm

Run on:          Thu Dec 18 08:52:32 1997   MasPar time 10.33 Seconds
                                            494.974 Million cell updates/sec

Tabular output not generated

Title:           >US-08-826-361A-6
Description:     (1-416) from US08826361A.pep
Perfect Score:   3108
Sequence:        1 MNYSIPSNVTNLEGGPGRQT..........QQSMPLANLLMLLSHVRHAR 416

Scoring table:   PAM 150
                 Gap 11

Searched.        101610 seqs, 12294212 residues

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database:        a-geneseq28
                   1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
                   8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
                   14:part14 15:part15 16:part16 17:part17 18:part18
                   19:part19 20:part20 21:part21

Statistics:      Mean 35.283,  Variance 145.558;  scale 0.242

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES

                 %
Result         Query
  No.  Score   Match Length DB  ID            Description           Pred. No.
 -----------------------------------------------------------------------------
    1   3099   99.7    485 21  W14724        Human oestrogen recep  0.00e+00
    2   2838   91.3    485 21  W14723        Rat oestrogen recepto  3.06e-277
    3   2820   90.7    485 21  W14725        Mouse oestrogen recep  2.29e-275
    4   1451   46.7    595  2  P70543        Human oestrogen recep  2.00e-133
    5   1422   45.8    334  2  P80925        Sequence of the human  1.94e-130
    6    736   23.7    433  2  P80931        Sequence of human est  2.26e-60
    7    725   23.3    521  2  P80930        Sequence of human est  2.91e-59
    8    614   19.8    777  2  P80919        Sequence of the prima  4.04e-48
    9    563   18.1     77 10  R53289        Human oestrogen recep  4.94e-43
   10    559   18.0    902  3  R12224        Rat androgen receptor  1.23e-42
   11    556   17.9    902  1  P91006        Rat androgen receptor  2.45e-42
   12    556   17.9    902  1  P93110        Rat androgen receptor  2.45e-42
   13    556   17.9    918  3  R12223        Human androgen recept  2.45e-42
   14    553   17.8    919  1  P90996        Human androgen recept  4.87e-42
   15    553   17.8    919 21  W14783        Androgen receptor.     4.87e-42
   16    552   17.8    919  1  P93109        Human androgen recept  6.13e-42
   17    546   17.6     72  1  P91994        Part of oestrogen rec  2.42e-41
   18    545   17.5    630  3  R12230        TrpE/androgen recepto  3.04e-41
   19    536   17.2     68 13  R71566        Partial human estroge  2.38e-40
   20    521   16.8    984  2  P80927        Sequence of the human  7.29e-39
```

```
   21    506   16.3    533  8  R39468        hRXR-beta1.            2.22e-37
   22    506   16.3    533 13  R72483        Human H-2RIIBP.        2.22e-37
   23    505   16.2    446 13  R72484        H-2RIIBP.              2.79e-37
   24    505   16.2    446  2  R11837        H2RIIBP nuclear hormo  2.79e-37
   25    501   16.1    448  8  R39467        mRXR-beta              6.93e-37
   26    501   16.1    510  8  R39469        hRXR-beta2.            6.93e-37
   27    493   15.9    462  3  R13795        Human RXR-alpha retin  4.27e-36
   28    494   15.9    467  8  R39470        mRXR-alpha.            3.40e-36
   29    494   15.9    467  3  R13799        Mouse RXR-alpha retin  3.40e-36
   30    471   15.2    598  9  R48631        Sequence of nuclear r  6.28e-34
   31    464   14.9    463  8  R39471        mRXR-gamma.            3.06e-33
   32    461   14.8    463  3  R13800        Mouse RXR-gamma retin  6.04e-33
   33    449   14.4    463 20  W00605        Human retinoid X rece  9.09e-32
   34    438   14.1    628 16  R92057        Apoptopic cerebral ne  1.09e-30
   35    420   13.5    447 20  W03326        LXR-alpha, orphan mem  6.23e-29
   36    412   13.3    448 12  R66764        Human RAR-beta.        3.75e-28
   37    412   13.3    448 17  R85315        Human hepatoma retino  3.75e-28
   38    411   13.2    440  7  R33744        XR2                    4.70e-28
   39    411   13.2    447  4  R20466        Mouse retinoic acid r  4.70e-28
   40    410   13.2    448 10  R55128        Retinoic acid recepto  5.88e-28
   41    411   13.2    454  2  R10182        Recombinant human gam  4.70e-28
   42    411   13.2    454  2  R10550        Human Retinoic Acid R  4.70e-28
   43    411   13.2    454 16  R84723        Human retinoic acid r  4.70e-28
   44    411   13.2    458  4  R20465        Mouse retinoic acid r  4.70e-28
   45    408   13.1    458  2  R10549        Skin-specific murine   9.20e-28

                              ALIGNMENTS

RESULT    1
ID   W14724 standard; Protein; 485 AA.
AC   W14724;
DT   08-JUN-1997  (first entry)
DE   Human oestrogen receptor beta (ER-beta).
KW   Orphan receptor; oestrogen receptor beta; ER-beta;
KW   nuclear receptor; prostate cancer; benign prostatic hyperplasia;
KW   ovary cancer; cardiovascular disease; osteoporosis;
KW   environment; pollutant.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   Domain          104..169
FT   /label= DNA-binding_domain
FT   Domain          260..457
FT   /label= Ligand-binding_domain
PN   WO9709348-A2.
PD   13-MAR-1997.
PF   09-SEP-1996; E03933.
PR   08-SEP-1995; GB-018272.
PR   15-MAR-1996; GB-005550.
PR   11-APR-1996; GB-007532.
PR   08-MAY-1996; GB-009576.
PA   (KARO-) KARO BIO AB
PI   Enmark E,  Gustafsson J,  Kuiper GG;
DR   WPI; 97-192842/17.
DR   N-PSDB; T62843.
PT   New isolated oestrogen receptor beta - used to develop prods. for
PT   treating e.g. cancers, CNS diseases, osteoporosis or cardiovascular
PT   disease
PS   Claim 1; Fig 13A; 45pp; English.
CC   A novel human orphan receptor (W14724) is related to the known
CC   oestrogen receptor ER-alpha, and has been designated ER-beta.
CC   It is an oestrogen receptor-related nuclear receptor.  The ER-beta
CC   amino acid sequence was deduced from a cDNA clone (T62843) isolated
CC   from a human testis cDNA library.  Rat, human and mouse ER-beta
CC   (W14723-25) can be used to isolate molecules for use in the
CC   treatment of cardiovascular diseases, central nervous system
CC   diseases, osteoporosis, prostate or ovarian cancer or benign
CC   prostatic hyperplasia and to test environmental chemicals for
CC   oestrogenic activity.
SQ   Sequence   485 AA;

  Query Match            99.7%;  Score 3099;  DB 21;  Length 485;
  Best Local Similarity  100.0%;  Pred. No. 0.00e+00;
  Matches   415;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Db          9 mnysipsnvtnleggpgrqttspnvlwptpghlsplvvhrqlshlyaepqkspwcearsl 68
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy          1 MNYSIPSNVTNLEGGPGRQTTSPNVLWPTPGHLSPLVVHRQLSHLYAEPQKSPWCEARSL 60

Db         69 ehtlpvnretlkrkvsgnrcaspvtgpgskrdahfcavcsdyasgyhygvwscegckaff 128
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy         61 EHTLPVNRETLKRKVSGNRCASPVTGPGSKRDAHFCAVCSDYASGYHYGVWSCEGCKAFF 120

Db        129 krsiqghndyicpatnqctidknrrksqacrlrkcyevgmvkcgsrrercgyrlvrrqr 188
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        121 KRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQR 180

Db        189 sadeqlhcagkakrsgghaprvrellldalspeqlvltlleaepphvlisrpsapfteas 248
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        181 SADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQLVLTLLEAEPPHVLISRPSAPFTEAS 240

Db        249 mmmsltkladkelvhmiswakkipgfvelslfdqvrllescwmevlmmglmwrsidhpgk 308
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        241 MMMSLTKLADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGK 300

Db        309 lifapdlvldrdegkcvegileifdmllattsrfrelklqhkeylcvkamillnssmypl 368
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        301 LIFAPDLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPL 360

Db        369 vtatqdadssrklahllnavtdalvwviaksgissqqqsmrlanllmllshvrha 423
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        361 VTATQDADSSRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHA 415
```

```
RESULT    2
ID   W14723 standard; Protein; 485 AA.
AC   W14723;
DT   08-JUN-1997  (first entry)
DE   Rat oestrogen receptor beta (ER-beta).
KW   Orphan receptor; oestrogen receptor beta; ER-beta;
KW   nuclear receptor; prostate cancer; benign prostatic hyperplasia;
KW   ovary cancer; cardiovascular disease; osteoporosis;
KW   environment; pollutant.
OS   Rattus sp.
FH   Key             Location/Qualifiers
FT   Domain          104..169
FT   /label= DNA-binding_domain
FT   Domain          259..457
FT   /label= Ligand-binding_domain
PN   WO9709348-A2.
PD   13-MAR-1997.
PF   09-SEP-1996; E03933
PR   08-SEP-1995; GB-018272.
PR   15-MAR-1996; GB-005550.
PR   11-APR-1996; GB-007532.
PR   08-MAY-1996; GB-009576
PA   (KARO-) KARO BIO AB.
PI   Enmark E,  Gustafsson J,  Kuiper GG;
DR   WPI; 97-192842/17
DR   N-PSDB; T62842.
PT   New isolated oestrogen receptor beta - used to develop prods. for
PT   treating e.g. cancers, CNS diseases, osteoporosis or cardiovascular
PT   disease
PS   Claim 1; Fig 1; 45pp; English.
CC   A novel rat orphan receptor (W14723) is related to the known
CC   oestrogen receptor ER-alpha, and has been designated ER-beta.
CC   It is an oestrogen receptor-related nuclear receptor.  The ER-beta
CC   amino acid sequence was deduced from a cDNA clone (T62842) isolated
CC   from a rat prostate cDNA library.  Rat, human and mouse ER-beta
CC   (W14723-25) can be used to isolate molecules for use in the
CC   treatment of cardiovascular diseases, central nervous system
CC   diseases, osteoporosis, prostate or ovarian cancer or benign
CC   prostatic hyperplasia and to test environmental chemicals for
CC   oestrogenic activity
SQ   Sequence   485 AA;
```

```
  Query Match             91.3%;  Score 2838;  DB 21;  Length 485;
  Best Local Similarity   89.1%;  Pred. No. 3.06e-277;
  Matches   369;  Conservative   29;  Mismatches   16;  Indels    0;  Gaps    0;

Db          9 mnysvpgstsnldggpvrlstspnvlwptsghlsplathcqssllyaepqkspwcearsl 68
              ||||:|:: :||:||| | :|||||||||:||||||  | | | ||||||||||||||||
Qy          1 MNYSIPSNVTNLEGGPGRQTTSPNVLWPTPGHLSPLVVHRQLSHLYAEPQKSPWCEARSL 60

Db         69 ehtlpvnretlkrklsgsscaspvtspnakrdahfcpvcsdyasgyhygvwscegckaff 128
              ||||||||||||||:||: ||||||:|  :|||||||:||||||||||||||||||||||
Qy         61 EHTLPVNRETLKRKVSGNRCASPVTGPGSKRDAHFCAVCSDYASGYHYGVWSCEGCKAFF 120

Db        129 krsiqghndyicpatnqctidknrrksqacrlrkcyevgmvkcgsrrercgyrivrrqr 188
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Qy        121 KRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQR 180

Db        189 ssseqvhclskakrngghaprvkelllstlspeqlvltlleaeppnvlvsrpsmpfteas 248
              |: ||:|| :||||:||||||:||||  :|||||||||||||||:||:||||| ||| ||
Qy        181 SADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQLVLTLLEAEPPHVLISRPSAPFTEAS 240

Db        249 mmmsltkladkelvhmigwakkipgfvelslldqvrllescwmevlmvglmwrsidhpgk 308
              |||||||||||||||||:|||||||||||||:|||||||||||||||:||||||||||||
Qy        241 MMMSLTKLADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGK 300

Db        309 lifapdlvldrdegkcvegileifdmllattsrfrelklqhkeylcvkamillnssmypl 368
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        301 LIFAPDLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPL 360

Db        369 asanqeaessrklthllnavtdalvwviaksgissqqqsvrlanllmllshvrh 422
               :| |:|:|||||:|||||||||||||||||||||||||:||||||||||||||
Qy        361 VTATQDADSSRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRH 414
```

```
RESULT    3
ID   W14725 standard; Protein; 485 AA.
AC   W14725;
DT   08-JUN-1997  (first entry)
DE   Mouse oestrogen receptor beta (ER-beta).
KW   Orphan receptor; oestrogen receptor beta; ER-beta;
KW   nuclear receptor; prostate cancer; benign prostatic hyperplasia;
KW   ovary cancer; cardiovascular disease; osteoporosis;
KW   environment; pollutant.
OS   Mus sp.
FH   Key             Location/Qualifiers
FT   Domain          104..169
FT   /label= DNA-binding_domain
FT   Domain          260..457
FT   /label= Ligand-binding_domain
PN   WO9709348-A2.
PD   13-MAR-1997.
PF   09-SEP-1996; E03933.
PR   08-SEP-1995; GB-018272.
PR   15-MAR-1996; GB-005550.
PR   11-APR-1996; GB-007532.
PR   08-MAY-1996; GB-009576.
PA   (KARO-) KARO BIO AB.
PI   Enmark E,  Gustafsson J,  Kuiper GG;
DR   WPI; 97-192842/17.
DR   N-PSDB; T62844.
PT   New isolated oestrogen receptor beta - used to develop prods. for
PT   treating e.g. cancers, CNS diseases, osteoporosis or cardiovascular
PT   disease
PS   Claim 1; Fig 14A; 45pp; English.
CC   A novel mouse orphan receptor (W14725) is related to the known
CC   oestrogen receptor ER-alpha, and has been designated ER-beta.
CC   It is an oestrogen receptor-related nuclear receptor.  The ER-beta
CC   amino acid sequence was deduced from a cDNA clone (T62844).
CC   Rat, human and mouse ER-beta (W14723-25) can be used to isolate
CC   molecules for use in the treatment of cardiovascular diseases,
CC   central nervous system diseases, osteoporosis, prostate or ovarian
CC   cancer or benign prostatic hyperplasia and to test environmental
CC   chemicals for oestrogenic activity.
SQ   Sequence   485 AA;
```

```
  Query Match             90.7%;  Score 2820;  DB 21;  Length 485;
  Best Local Similarity   88.6%;  Pred. No. 2.29e-275;
  Matches  367;  Conservative   31;  Mismatches   16;  Indels    0;  Gaps    0;

Db        9 mnysvpsstgnleggpvrqtaspnvlwptsghlsplathrqssllyaepqkspwpearsl 68
            ||||:||:  |||||| |||:||||||||:||||||  ||| | |||||||||| |||||
Qy        1 MNYSIPSNVTNLEGGPGRQTTSPNVLWPTPGHLSPLVVHRQLSHLYAEPQKSPWCEARSL 60

Db       69 ehtlpvnretlkrklqqsgqaspvtspstkrdahfcavcsdyasgyhygvwscegckaff 128
            ||||||||||||||:: :  |||||:|::|||||||||||||||||||||||||||||||
Qy       61 EHTLPVNRETLKRKVSGNRCASPVTGPGSKRDAHFCAVCSDYASGYHYGVWSCEGCKAFF 120

Db      129 krsiqghndyicpatnqctidknrrkscqacrlrkcyevgmvkcgsrrercgyrivrrqr 188
            |||||||||||||||||||||||||||:||||||||||||||||||||||||||:|||||
Qy      121 KRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQR 180

Db      189 saseqvhclnkakrtsghtprvkelllnslspeqlvltlleaeppnvlvsrpsmpfteas 248
            || ||:||  ||||::||.|||:||||..||||||||||||||||:||:|||| ||||||
Qy      181 SADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQLVLTLLEAEPPHVLISRPSAPFTEAS 240

Db      249 mmmsltkladkelvhmigwakkipgfvelslldqvrllescwmevlmvglmwrsidhpgk 308
            |||||||||||||||||:|||||||||||||:||||||||||||||||:|||||||||||
Qy      241 MMMSLTKLADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGK 300

Db      309 lifapdlvldrdegkcvegileifdmllattarfrelklqhkeylcvkamillnssmyhl 368
            |||||||||||||||||||||||||||||||:|||||||||||||||||||||||||| |
Qy      301 LIFAPDLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPL 360

Db      369 atasqeaessrklthllnavtdalvwvisksrissqqqsvrlanllmllshvrh 422
             ||:|:|:|||||:||||||||||||||:|| ||||||||:|||||||||||||
Qy      361 VTATQDADSSRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRH 414


RESULT    4
ID   P70543 standard; protein; 595 AA
AC   P70543;
DT   24-FEB-1991  (first entry)
DE   Human oestrogen receptor protein.
KW   Human oestrogen receptor protein; expression system; assay;
OS   Homo sapiens
PN   WO8705049-A.
PD   27-AUG-1987
PF   18-FEB-1987; U00341.
PR   20-FEB-1986; US-833829
PA   (CALB-) California Biotechnology Inc.
PI   Shine J;
DR   WPI; 87-250213/35.
DR   N-PSDB; N70880.
PT   Expression system for vertebrate steroid receptor protein - comprising
PT   DNA sequence encoding the protein linked to control sequences in
PT   eucaryotic hosts.
PS   Disclosure; Fig. 1-1 - 1-2; 26pp; English.
CC   steroid receptor protein;
CC   The protein is expressed in an expression system from cDNA operably
CC   linked to control sequences compatible with eukaryotic host cells
CC   This method allows expression under conditions which favour appropriate
CC   post-translational processing.  It also produces large amts. of purified
CC   protein useful in the design of agonist and antagonist cpds. for study
CC   of the mechanism of action of the steroid binding proteins in general.
CC   and for use in diagnostic assays for the proteins or antibodies to
CC   them.  These assays are important in eg the diagnosis of tumour
CC   sensitivities to steroid metabolism.  Suitable host cells are VERO,
CC   HeLa and CHO cells
SQ   Sequence   595 AA;

  Query Match             46.7%;  Score 1451;  DB 2;  Length 595;
  Best Local Similarity   57.1%;  Pred. No. 2.00e-133;
  Matches  194;  Conservative   84;  Mismatches   47;  Indels   15;  Gaps   11;

Db      178 saketrycavcndyasgyhygvwscegckaffkrsiqghndymcpatnqctidknrrksc 237
            | ::  :||||.||||||||||||||||||||||||||||||:|||||||||||||||||
Qy       89 SKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSC 148

Db      238 qacrlrkcyevgmmkggirkdrrggrmlkhkrqrddgegrgevgsagdmraanlwpsplm 297
            |||||||||||||.| | |..| | |:..  |||   | .   .. : .|.:    .|
Qy      149 QACRLRKCYEVGMVKCGSRRERCGYRLVR--RQRSADE-QLHCAGKA-KRSGG--HAP-R 201

Db      298 ikrskknslalsltadqmvsalldaepp-ilyseydptrpfseasmmglltnladrelvh 356
            : |     |   |: :|:| .||.|||| :| |   |. ||.|||||  ||.|||:||||
Qy      202 V-RELL--LD-ALSPEQLVLTLLEAEPPHVLISR--PSAPFTEASMMMSLTKLADKELVH 255

Db      357 minwakrvpgfvdltlhdqvhllecawleilmiglvwrsmehpvkllfapnllldrnqgk 416
            ||:|||::||||:|:| |||.|||  | |:||:||:|||::|| ||:|||:|:|||::||
Qy      256 MISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGK 315

Db      417 cvegmveifdmllatssrfrmmnlqgeefvclksiillnsgvytflsstlksleekdhih 476
            ||||::|||||||||:|||| :.||  |.:|:|.:|||||.:|  :..|         :
Qy      316 CVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSRKLA 374

Db      477 rvldkitdtlihlmakagltlqqqhqrlaqlllilshirh 516
            .:|. :||.|: .:||.|:. |||  ||| ||::|||:||
Qy      375 HLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRH 414


RESULT    5
ID   P80925 standard; protein; 334 AA.
AC   P80925;
DT   29-DEC-1990  (first entry)
DE   Sequence of the human oestrogen receptor
KW   Hormone receptor; hormone-binding, transcription activation;.
OS   Homo sapiens.
PN   WO8803168-A.
PD   05-MAY-1988.
PF   23-OCT-1987, U02782.
PR   20-OCT-1987; US-108471.
PA   (SALK) Salk Inst for Biol Stud.
PI   Evans RM, Weinberger CA, Hollenberg SM, Giguere V;
DR   WPI; 88-133242/19.
PT   Recombinant DNA encoding hormone recpetors -
PT   comprising glucocorticoid, mineralocorticoid, thyroid hormone
PT   and novel hormone receptors
PS   Example, Fig III-2, 243pp; English.
CC   DNAs encoding hormone receptors and the hormone receptors themselves are
CC   claimed. The DNA can be used to make the hormone receptor proteins and
CC   functional modified forms in quantities not previously possible. The
CC   receptor proteins can be used to screen cpds. for receptor-agonist or
CC   receptor-antagonist activity. They can also be used in diagnostic assays.
SQ   Sequence   334 AA;

  Query Match             45.8%;  Score 1422;  DB 2;  Length 334;
  Best Local Similarity   57.5%;  Pred. No. 1.94e-130;
  Matches  192;  Conservative   80;  Mismatches   46;  Indels   16;  Gaps   12;

Db        1 ycavcndyasgyhygvwscegckaffkrs-qghndymcpatnqctidknrrkscqacrlr 59
            .||||:||||||||||||||||||||||| ||||||:|||||||||||||||||||||||
Qy       95 FCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLR 154

Db       60 kcyrvgmmkggirkdrrggrmlkhkrqrddgegrgevgsagdmraanlwpsplmikrskk 119
            ||| |||.| | |..| | |...  |||    | :    :: : :|::    :|   : |
Qy      155 KCYEVGMVKCGSRRERCGYRLVR--RQRSADE-QLHCAGKA-KRSGG--HAP-RV-RELL 206

Db      120 nslalsltadqmvsalldaepp-ilyseydptrpfseasmmglltnladrelvhminwak 178
              |  .|..:|:| .||:|||| :| |   |. ||.|||||  ||.|||:||||||:|||
Qy      207 --LD-ALSPEQLVLTLLEAEPPHVLISR--PSAPFTEASMMMSLTKLADKELVHMISWAK 261

Db      179 rvpgfvdltlhdqvhllecawleilmiglvwrsmehpvkllfapnllldrnqgkcvegmv 238
            ::||||:|:| |||.|||  | |:||:||:|||::|| ||:|||:|:|||::||||||::
Qy      262 KIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGIL 321

Db      239 eifdmllatssrfrmmnlqgeefvclksiillnsgvytflsstlksleekdhihrvldki 298
            |||||||||:|||| :.||  |.:|:|.:|||||.:|  :..|         :  :|: :
Qy      322 EIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSRKLAHLLNAV 380

Db      299 tdtlihlmakagltlqqqhqrlaqlllilshirh 332
            ||.|: ..||.|:. |||  ||| ||:.|||.||
Qy    381 TDALVWVIAKSGISSQQQSMRLANLLMLLSHVRH 414


RESULT    6
ID   P80931 standard; protein; 433 AA.
AC   P80931;
DT   29-DEC-1990  (first entry)
DE   Sequence of human estrogen-related receptor protein (hERR2)
KW   Hormone receptor; hormone-binding, transcription activation
OS   Homo sapiens.
PN   WO8803168-A.
PD   05-MAY-1988.
PF   23-OCT-1987; U02782.
PR   20-OCT-1987; US-108471.
PA   (SALK) Salk Inst for Biol Stud.
PI   Evans RM, Weinberger CA, Hollenberg SM, Giguere V;
DR   WPI; 88-133242/19.
DR   N-PSDB; N80922.
PT   Recombinant DNA encoding hormone recpetors -
PT   comprising glucocorticoid, mineralocorticoid, thyroid hormone
PT   and novel hormone receptors
PS   Claim 17; Fig V-2(B)-1 and -2; 243pp; English.
CC   DNAs encoding hormone receptors and the hormone receptors themselves are
CC   claimed. The DNA can be used to make the hormone receptor proteins and
CC   functional modified forms in quantities not previously possible. The
CC   receptor proteins can be used to screen cpds. for receptor-agonist or
CC   receptor-antagonist activity. They can also be used in diagnostic assays.
CC   Also claimed is pure DNA encoding oestrogen-related receptors hERR1 and
CC   hERR2 and hERR1 and hERR2. The new hERR1 and hERR2 receptors will provide
CC   the basis for development of an assay system that will lead to the
CC   identification of novel hormones.
SQ   Sequence   433 AA;

  Query Match             23.7%;  Score 736;  DB 2;  Length 433;
  Best Local Similarity   43.6%;  Pred. No. 2.26e-60;
  Matches   116;  Conservative   55;  Mismatches  84;  Indels  11;  Gaps   9;

Db    101 rlclveqdiasqyhyqvasceackaftkrtiqqnieyscpatneceitkrrrkscqacrf 160
            ::| ||:| |||||||| |||:|||||||||:|||: :| ||||||:| | | |||||||||:
Qy     94 HFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRL 153

Db    161 mkelkvqmlkeqvrldrv  r qqrqky krrldsenspylslqisppakkpltkivs-- 214
             ||   |||:| | | :|    |    ||:     :|        :  |    :|  :  |    :|
Qy    154 RKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPE 213

Db    215 yl l vaepdklyam ppddvpegdikalttlcdladrelvfliswakhipgfsnltlgd 271
             | |  : |::      : :      |    :      : :|  |||:||| :||||| |||| :|:| |
Qy    214 QLVLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 273

Db    272 qmsllqsawmellilqivyrslpydgklvyaedyimdeehsrlv-gllelyraillqlvrr 330
            |: ||:| |||:|::|:: ||:      || :| | ::| : :: | |:||:: :|      |
Qy    274 QVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSR 333

Db    331 ykklkvekeefvmlkalalansdsmy 356
            :: ||::  |:: :||: | || |||
Qy    334 FRELKLQHKEYLCVKAMILLNS-SMY 358


RESULT    7
ID   P80930 standard; protein; 521 AA.
AC   P80930;
DT   10-MAR-1993  (revised)
DT   29-DEC-1990  (first entry)
DE   Sequence of human estrogen-related receptor protein (hERR1)
KW   Hormone receptor; hormone-binding; transcription activation.
OS   Homo sapiens.
PN   WO8803168-A.
PD   05-MAY-1988.
PF   23-OCT-1987; U02782.
PR   20-OCT-1987; US-108471.
PA   (SALK) Salk Inst for Biol Stud.
PI   Evans RM, Weinberger CA, Hollenberg SM, Giguere V;
DR   WPI; 88-133242/19.
DR   N-PSDB; N80921.
PT   Recombinant DNA encoding hormone recpetors -
PT   comprising glucocorticoid, mineralocorticoid, thyroid hormone
PT   and novel hormone receptors
PS   Claim 17; Fig V-1(B)-1 and -2; 243pp; English.
CC   DNAs encoding hormone receptors and the hormone receptors themselves are
CC   claimed. The DNA can be used to make the hormone receptor proteins and
CC   functional modified forms in quantities not previously possible. The
CC   receptor proteins can be used to screen cpds. for receptor-agonist or
CC   receptor-antagonist activity. They can also be used in diagnostic assays.
CC   Also claimed is pure DNA encoding oestrogen-related receptors hERR1 and
CC   hERR2 and hERR1 and hERR2. The new hERR1 and hERR2 receptors will provide
CC   the basis for development of an assay system that will lead to the
CC   identification of novel hormones.
SQ   Sequence   521 AA;

  Query Match             23.3%;  Score 725;  DB 2;  Length 521;
  Best Local Similarity   40.1%;  Pred. No. 2.91e-59;
  Matches   130;  Conservative   67;  Mismatches 117;  Indels  10;  Gaps   7;

Db    174 rlclvcgdvasgyhygvasceackaffkrtiqgsieyscpasneceitkrrrkacqacrf 233
            .:| ||:| |||||||| |||:|||||||||:|||  :| |||:|:| | | |||:|||||:
Qy     94 HFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRL 153

Db    234 tkclrvgmlkegvrldrvrggrqkykrrpevdplpfpg-p-fpagplavaggprktaapv 291
             ||  |||.| | | .|  | |     .:|.   . |   .| :   ::|  :         | :
Qy    154 RKCYEVGMVKCGSRRERC-GYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSP 212

Db    292 nalvshllvvepeklyampdpagpdghlpavatlcdlfdrelvvtiswaksipgfsslsl 351
            : ||  ||  ||  :  :: |::|     : : :|  | |:|:|  ||||| |||| |||
Qy    213 EQLVLTLLEAEPPHVL-ISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSL 271

Db    352 sdqmsvlqsvwmevlvlgvaqrslplqdelafaedlvld-eegaraaglgelgaallqlv 410
             ||: :|:| |||||::|: ||:       | || ||||| :||      |: |:     ||
Qy    272 FDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATT 331

Db    411 rrlqalrlereeyvllkalalansdsvhiedeprlwsscekllh--eal---leyeaqra 465
             |:: |:|:: ||: :||: | ||      :     :     |  || |  :|:    |    :::
Qy    332 SRFRELKLQHKEYLCVKAMILLNSSMYPLVIATQDADSSRKLAHLLNAVTDALVWVIAKS 391

Db    466 gpgggaerrragrllltlpllrqt 489
            | ::   :   | : ||: |: :|::
Qy    392 GISSQQQSMRLANLLMLLSHVRHA 415


RESULT    8
ID   P80919 standard; protein; 777 AA.
AC   P80919;
DT   29-DEC-1990  (first entry)
DE   Sequence of the primary protein sequence of human glucocorticoid
DE   receptor (hGR)
KW   Hormone receptor; hormone-binding; transcription activation.
OS   Homo sapiens.
PN   WO8803168-A.
PD   05-MAY-1988.
PF   23-OCT-1987; U02782.
PR   20-OCT-1987; US-108471.
PA   (SALK) Salk Inst for Biol Stud.
PI   Evans RM, Weinberger CA, Hollenberg SM, Giguere V;
DR   WPI; 88-133242/19.
DR   N-PSDB; N80916.
PT   Recombinant DNA encoding hormone recpetors -
PT   comprising glucocorticoid, mineralocorticoid, thyroid hormone
PT   and novel hormone receptors
PS   Claim 24; Fig I-2(1) to (2); 243pp; English.
CC   DNAs encoding hormone receptors and the hormone receptors themselves are
CC   claimed. The DNA can be used to make the hormone receptor proteins and
CC   functional modified forms in quantities not previously possible. The
CC   receptor proteins can be used to screen cpds. for receptor-agonist or
CC   receptor-antagonist activity. They can also be used in diagnostic assays.
SQ   Sequence   777 AA;

  Query Match             19.8%;  Score 611;  DB 2;  Length 777;
  Best Local Similarity  33.4%;  Pred. No. 4.04e-48;
  Matches   97;  Conservative   84;  Mismatches  93;  Indels  16;  Gaps  15;

Db     397 smrpdvsspssssstattgp-p-pklclvcsdeasgchygvltcgsckvffkravegqhn 454
           :::  ||.    .|. .  |.    . :| |||| ||| |||| .| :|| ||||:..|...
Qy      70 TLKRKVSGNRCASPVIGPGSKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHND 129

Db     455 ylcagrndciidkirrkncpacryrkclqagmnl-earktk-k-kikgiqqattg-v-sq 509
           |:|:: |:| ||| |||.| ||| ||| . ||      :|. :    .     |...  .
Qy     130 YICPATNQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCA 189

Db     510 etsenpgnkt-ivpatlpq-ltpt-lv-slleviepevlyagydssvpdstwrimttlnm 565
                .  .|   .   |    | . |:|  || .|||   | ||  .   ::  ...  .| :| :
Qy     190 GKAKRSGGHAPRVRELLLDALSPEQLVLTLLEAEPPHVLISRPSAPFTEAS--MMMSLTK 247

Db     566 lggrqviaavkwakaipgfrnlhlddqmtllqyswmflmafalgwrsyrqssanllcfap 625
           |. .:..   . ||| |||| .| | ||. ||: || ..   :| |||  ::: :|: |||
Qy     249 LADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPG-KLI-FAP 305

Db     626 dliin-eqrmtlpcmydqckhmlyvsselhrlqvsyeeylcmktllllss 674
           ||:::: :: : :  .  .       |    |   | .   ||||:|.:.:||:|
Qy     306 DLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNS 355

RESULT    9
ID   R53269 standard; protein, 77 AA.
AC   R53289;
DT   30-DEC-1994  (first entry)
DE   Human oestrogen receptor DNA-binding domain (185-261).
KW   Human; oestrogen receptor; DNA-binding domain; hER;
KW   monoclonal antibody; diagnosis; antiserum; cancer; hapten;
KW   immunogen.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   Peptide         17..31
FT   /label= peptide_1
FT   /note= "hER 201-215"
FT   Peptide         46..61
FT   /label= peptide_2
FT   /note= "hER 231-245"
FT   Peptide         63..77
FT   /label= peptide_3
FT   /note= "hER 247-261"
PN   US5312752-A.
PD   17-MAY-1994.
PF   31-JUL-1989; 388091.
PR   31-JUL-1989; US-388091
PR   01-NOV-1991; US-784626.
PA   (UYBO-) UNIV BOSTON.
PI   Traish AM,  Wotiz HH;
DR   WPI; 94-159123/19.
PT   Monoclonal antibody to human oestrogen receptor protein -
PT   specific for activated forms, useful for diagnosis, esp. in
PT   breast cancer assessment
PS   Disclosure; Fig 1; 41pp; English.
CC   The DNA-binding region of human oestrogen receptor (hER) is
CC   encompassed by the amino acids 180-251 in the native protein.
CC   Three peptides chosen as haptens and immunogens to obtain specific
CC   antisera, are indicated in the features table
SQ   Sequence   77 AA;

  Query Match             18.1%;  Score 563;  DB 10;  Length 77;
  Best Local Similarity   89.6%;  Pred. No. 4.94e-43;
  Matches   69;  Conservative    5;  Mismatches   3;  Indels   0;  Gaps   0;

Db       1 cavcndyasgyhygvwscegckaffkrsiqghndymcpatnqctidknrrkscqacrlrk 60
           ||||:||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Qy      96 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 155

Db      61 cyevgmmkggirkdrrg 77
           ||||||:| | |::| |
Qy     156 CYEVGMVKCGSRRERCG 172
```

```
RESULT   10
ID   R12224 standard; Protein; 902 AA.
AC   R12224;
DT   20-AUG-1991  (first entry)
DE   Rat androgen receptor
KW   rAR; DNA-binding protein; steroid hormone.
OS   Rattus rattus.
FH   Key             Location/Qualifiers
FT   Domain          540..610
FT   /label= DNA-binding domain
FT   /note= "cysteine-rich"
PN   WO9107423-A.
PD   30-MAY-1991.
PF   19-OCT-1990; U06015.
PR   17-NOV-1989; US-438775.
PA   (ARCH-) ARCH DEV CORP.
PI   Liao S, Chang C;
DR   WPI; 91-178048/24.
DR   N-PSDB; Q12002
PT   Androgen receptor and TR2 DNA binding proteins - DNA sequences
PT   and antibodies for detection and quantification methods
PS   Claim 25; Fig 3; 79pp; English.
CC   This sequence was deduced from a cDNA clone isolated by screening
CC   a rat ventral prostate lambda gt11 library in E.coli Y1090. The
CC   sequence is very similar to that of human AR and in
CC   the DNA-binding domain it is identical to that of hAR DNA-binding
CC   domain. Homology comparisons with other known steroid receptors
CC   indicate that rAR is more closely related to glucocorticoid,
CC   mineralo-corticoid and progesterone receptors than to v-erb-A or to
CC   receptors for oestrogen, vitamin D and thyroid hormones.
SQ   Sequence   902 AA;

  Query Match             18.0%;  Score 559;  DB 3;  Length 902;
  Best Local Similarity   33.5%;  Pred. No. 1.23e-42;
  Matches   88;  Conservative   72;  Mismatches  92;  Indels  11;  Gaps   8;

Db     542 clicgdeasgchygaltcgsckvffkraaegkqkylcasrndctidkfrrkncpscrlrk 601
           | :|:| ||| |||   :|    || ||||.   |      | |.   |:||||| |||:| :|||||
Qy      96 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 155

Db     602 cyeagmtl-garklk---klgnlklqeegenssagsptedpsqkmtvshi--egyecqpi 655
           ||| ||   |:|. .    .|           : .    ||  :     ..    |  :  ::   : :
Qy     156 CYEVGMVKCGSRREPCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 215

Db     656 flnvlcaicpgvvcaghdnnqpdsfaallsslnelgerqlvhvvkwakalpgfrnlhvdd 715
             | :||| | |:   :. .  | : |. . ||   | :.:|||:. ||| :||| :| : |
Qy     216 VLTLLEAEPPHVL-ISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 273

Db     716 qmavieyswmglmvfamgwrsftnvnsrmlyfapdlvfn-eyrmhksrmysqcvrmrhls 774
           |. .:|  || :.   :: |||    .    ....|||||||:: :   :      :    :   :
Qy     274 QVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDMLLATT 331

Db     775 qefgwlqitpqeflcmkalllfs 797
             |  | :    |.||.||:.|..
Qy     332 SRFRELKLQHKEYLCVKAMILLN 354

RESULT   11
ID   P91006 standard, protein; 902 BP.
AC   P91006;
DT   28-FEB-1990  (first entry)
DE   Rat androgen receptor DNA clone.
KW   Androgen receptor; TR2 polypeptide;
OS   Rat.
FH   Key             Location/Qualifiers
FT   Region          1..902
FT   /*tag=  a
FT   /product=98 kD polypeptide
FT   Region          170..902
FT   /*tag=  b
FT   /product=79 kD polypeptide
PN   WO8909223-A.
```

```
PD   05-OCT-1989.
PF   24-MAR-1989; U01238.
PR   30-MAR-1988; US-176107.
PA   (ARCH-) Arch Development Corp.
PI   Liao S, Chang C;
DR   WPI; 89-309501/42.
DR   N-PSDB; N91578.
PT   New DNA encoding new androgen receptor and TR2 polypeptide(s) - able
PT   to bind DNA, and derived antibodies, useful for receptor assay and
PT   purification.
PS   Claim 8; fig. 3; 60pp; English.
CC   The protein is used to raise antibodies for receptor assays and for
CC   affinity purification.
CC   The 98 kD product starts at the first Met codon, the 79 kD product
CC   starts from the second
SQ   Sequence   902 AA;

  Query Match             17.9%;  Score 556;  DB 1;  Length 902;
  Best Local Similarity  33.1%;  Pred. No. 2.45e-42;
  Matches   87;  Conservative   73;  Mismatches  92;  Indels  11;  Gaps   8;

Db     542 clicgdeasgchygaltcgsckvffkraaegkqkylcasrndctidkfrrkncpscrlrk 601
           | :|:| ||| |||  :| :|| ||||: :|   |:|:: |:||||| |||:| :|||||
Qy      96 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 155

Db     602 cyeagmtl-garklk---klgnlklqeegenssaqsptedpsqkmtvshi--egyecqpi 655
           ||| ||   |:|: :   :|         : :   || :    ::   |  :  ::   : :
Qy     156 CYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 215

Db     656 flnvlealepgvvcaghdnnqpdsfaallsslnelgerqlvhvvkwakalpgfrnlhvdd 715
            | :|||  | |:  :: :  | : |::: ||  |::::|||:: ||| :||| :| : |
Qy     216 VLTLLEAEPPHVL-ISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 273

Db     716 qmaviqyswmglmvfamgwrsftnvnsrmlyfapdlvfn-eyrmhksrmysqcvrmrhls 774
           |: ::: || ::: :: |||: :  :::: ||||||:: :  :     :         :  :
Qy     274 QVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDMLLATT 331

Db     775 qefgwlqitpqeflcmkalllfs 797
             |  | :   |:||:||::|::
Qy     332 SRFRELKLQHKEYLCVKAMILLN 354

RESULT   12
ID   P93110 standard; protein; 902 AA.
AC   P93110;
DT   19-MAR-1990  (first entry)
DE   Rat androgen receptor.
KW   Rat androgen receptor; monoclonal antibody; ploy-clonal antibody; cancer
OS   Rattus rattus.
PN   WO8909791-A.
PD   19-OCT-1989.
PF   13-APR-1989; U01548.
PR   14-APR-1988; US-182646.
PA   (UYNC-) University of North Carolina.
PI   French FS, Wilson EM, Joseph DR, Lubahn DB;
DR   WPI; 89-324206/44.
DR   N-PSDB; N91773.
PT   DNA encoding androgen receptor protein - useful for transforming
PT   eukaryotic hosts for protein expression and subsequent antibody prodn.
PS   Disclosure; Fig. 5; 41pp; English.
CC   Androgen receptor protein (AR) is used to produce mono- or poly-clonal
CC   antibodies.  These are used for the detection and quantification of AR in
CC   the presence of endogenous androgen, as androgen will not interfere with
CC   binding.  They may be used in assays to determine and quantify cellular
CC   distribution of AR in tumour tissue, and are esp. useful for evaluating
CC   prostate cancers to determine responsiveness to androgen withdrawal
CC   therapy.
SQ   Sequence   902 AA;

  Query Match             17.9%;  Score 556;  DB 1;  Length 902;
  Best Local Similarity  33.1%;  Pred. No. 2.45e-42;
  Matches   87;  Conservative   73;  Mismatches  92;  Indels  11;  Gaps   8;
```

```
Db     542 clicgdeasgchygaltcgsckvffkraaegkqkylcasrndctidkfrrkncpscrlrk 601
           | :|:| ||| |||  :| :|| ||||: :|   |:|:: |:||||| |||:| :|||||
Qy      96 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 155

Db     602 cyeagmtl-garklk---klgnlklqeegenssaqsptedpsqkmtvshi--egyecqpi 655
           ||| ||   |:|: :   :|         : :   || :    ::   |  :  ::   : :
Qy     156 CYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 215

Db     656 flnvlealepgvvcaghdnnqpdsfaallsslnelgerqlvhvvkwakalpgfrnlhvdd 715
            | :|||  | |:  :: :  | : |::: ||  |::::|||:: ||| :||| :| : |
Qy     216 VLTLLEAEPPHVL-ISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 273

Db     716 qmaviqyswmglmvfamgwrsftnvnsrmlyfapdlvfn-eyrmhksrmysqcvrmrhls 774
           |       ||        |||      ::: ||||||:: :  :     :         :  :
Qy     274 QVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDMLLATT 331

Db     775 qefgwlqitpqeflcmkalllfs 797
             |  | :   |:||:||::|::
Qy     332 SRFRELKLQHKEYLCVKAMILLN 354

RESULT   13
ID   R12223 standard; Protein; 918 AA.
AC   R12223;
DT   20-AUG-1991  (first entry)
DE   Human androgen receptor.
KW   hAR; DNA-binding protein; steroid hormone.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   Domain          556..626
FT   /label= DNA-binding domain
FT   /note= "cysteine-rich"
PN   WO9107423-A.
PD   30-MAY-1991.
PF   19-OCT-1990; U06015.
PR   17-NOV-1989; US-438775.
PA   (ARCH-) ARCH DEV CORP.
PI   Liao S, Chang C;
DR   WPI; 91-178048/24.
DR   N-PSDB; Q12001.
PT   Androgen receptor and TR2 DNA binding proteins - DNA sequences
PT   and antibodies for detection and quantification methods
PS   Claim 25; Fig 3; 79pp; English.
CC   This sequence was deduced from a cDNA clone isolated by screening
CC   commercially available human testis and prostate lambda gt11 cDNA
CC   libraries. The sequence is very similar to that of rat AR and in
CC   the DNA-binding domain it is identical to that of rAR DNA-binding
CC   domain. Homology comparisons with other known steroid receptors
CC   indicate that hAR is more closely related to glucocorticoid,
CC   mineralo-corticoid and progesterone receptors than to v-erb-A or to
CC   receptors for oestrogen, vitamin D and thyroid hormones.
SQ   Sequence   918 AA;

  Query Match             17.9%;  Score 556;  DB 3;  Length 918;
  Best Local Similarity  32.7%;  Pred. No. 2.45e-42;
  Matches   86;  Conservative   76;  Mismatches  90;  Indels  11;  Gaps   9;

Db     558 clicgdeasgchygaltcgsckvffkraaegkqkylcasrndctidkfrrkncpscrlrk 617
           | :|:| ||| |||  :| :|| ||||: :|   |:|:: |:||||| |||:| :|||||
Qy      96 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 155

Db     618 cyeagmtl-garklk---klgnlklqeeg-eassttspteettqkltvshiegyecqpi 671
           ||| ||   |:|: :   :|     :  :|     :: :  :    : ::   :::   : :
Qy     156 CYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 215

Db     672 flnvlealepgvvcaghdnnqpdsfaallsslnelgerqlvhvvkwakalpgfrnlhvdd 731
            | :|||  | |:  :: :  | : |::: ||  |::::|||:: ||| :||| :| : |
Qy     216 VLTLLEAEPPHVL-ISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 273

Db     732 qmaviqyswmglmvfamgwrsftnvnsrmlyfapdlvfn-eyrmhksrmysqcvrmrhls 790
           |: ::| || ::: :: |||: :  :::: ||||||:: :  :     :         :  :
Qy     274 QVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDMLLATT 331
```

```
Db    791 qefgwlqitpqeflcmkalllfs 813
            |  | :   |:||:||::|::
Qy    332 SRFRELKLQHKEYLCVKAMILLN 354


RESULT   14
ID   P90996 standard; protein; 919 AA.
AC   P90996;
DT   28-FEB-1990  (first entry)
DE   Human androgen receptor DNA clone.
KW   Androgen receptor; TR2 polypeptide;
OS   Homo sapiens.
FH   Key              Location/Qualifiers
FT   Region           1..919
FT   /*tag=  a
FT   /product=99 kD polypeptide
FT   Region            185..919
FT   /*tag=  b
FT   /product=79 kD polypeptide
PN   WO8909223-A.
PD   05-OCT-1989.
PF   24-MAR-1989; U01238.
PR   30-MAR-1988; US-176107.
PA   (ARCH-) Arch Development Corp.
PI   Liao S, Chang C;
DR   WPI; 89-309501/42.
DR   N-PSDB; N91577.
PT   New DNA encoding new androgen receptor and TR2 polypeptide(s) - able
PT   to bind DNA, and derived antibodies, useful for receptor assay and
PT   purification.
PS   Claim 8; fig. 3; 60pp; English.
CC   The protein is used to raise antibodies for receptor assays and for
CC   affinity purification.
CC   The 99 kD product starts at the first Met codon, the 79 kD product
CC   starts from the second.
SQ   Sequence   919 AA;

  Query Match            17.8%;  Score 553;  DB 1;  Length 919;
  Best Local Similarity  32.3%;  Pred. No. 4.87e-42;
  Matches   85;  Conservative   77;  Mismatches  90;  Indels  11;  Gaps   9;

Db    559 clicgdeasgchygaltcgsckvffkraaegkqkylcasrndctidkfrrkncpscrlrk 618
          | :|:| ||| |||  .| .|| ||||. :|   |.|:: |:||||| |||:| :|||||
Qy     96 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 155

Db    619 cyeagmtl-garklk---klgn-lklqeeg-eassttspteettqkltvshiegyecqpi 672
          |||  ||   |:|: :   :|    :  :|    ::. . .   . :     ::.  .  :
Qy    156 CYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 215

Db    673 flnvleaiepgvvcaghdnnqpdsfaallsslnelgerqlvhvvkwakalpgfrnlhvdd 732
           |.:|||  | |:  :  :  | . |... ||  |.:::|||:: ||| :||| :| : |
Qy    216 VLTLLEAEPPHVL-ISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 273

Db    733 qmaviqyswmglmvfamgwrsftnvnsrmlyfapdlvfn-eyrmhksrmysqcvrmrhls 791
          |: :::  ||  ::  .: |||  .  .::: ||||||.: :       :      :   :
Qy    274 QVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDMLLATT 331

Db    792 qefgwlqitpqeflcmkalllfs 814
            |  | :   |:||:||::|::
Qy    332 SRFRELKLQHKEYLCVKAMILLN 354


RESULT   15
ID   W14783 standard; Protein; 919 AA.
AC   W14783;
DT   22-JUN-1997  (first entry)
DE   Androgen receptor.
KW   Androgen receptor; acidic fibroblast growth factor; aFGF;
KW   antisense; benign prostatic hyperplasia; prostate cancer; therapy.
OS   Homo sapiens.
PN   WO9711170-A1.
PD   27-MAR-1997.
PF   20-SEP-1996; U15081.
PR   20-SEP-1995; US-004018.
PA   (WORC-) WORCESTER FOUND BIOMEDICAL RES.
PI   Zamecnik PA;
DR   WPI; 97-202879/18.
DR   N-PSDB; T63407.
PT   Oligonucleotide(s) antisense to human androgen receptor and acidic
PT   FGF genes - used to inhibit gene expression, for the treatment of
PT   benign prostatic hyperplasia
PS   Disclosure; Page 22-28; 51pp; English.
CC   Human androgen receptor (W14783) binds testosterone and, acting
CC   at the transcriptional level, regulates the growth of normal
CC   prostatic cells.  Antisense oligonucleotides (see also T63200,
CC   T63404-05) based on an androgen receptor cDNA clone (see also
CC   T63407) can be used to prevent androgen receptor gene expression,
CC   thereby inhibiting the growth or survival of prostatic cells for
CC   the treatment of benign prostatic hyperplasia and prostate cancer.
SQ   Sequence   919 AA;

  Query Match            17.8%;  Score 553;  DB 21;  Length 919;
  Best Local Similarity  32.3%;  Pred. No. 4.87e-42;
  Matches   85;  Conservative   77;  Mismatches  90;  Indels  11;  Gaps   9;

Db    559 clicgdeasgchygaltcgsckvffkraaegkqkylcasrndctidkfrrkncpscrlrk 618
          | :|:| ||| |||  .| .|| ||||. :|   |.|:: |:||||| |||:| :|||||
Qy     96 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 155

Db    619 cyeagmtl-garklk---klgn-lklqeeg-eassttspteettqkltvshiegyecqpi 672
          |||  ||   |:|: :   :|    :  :|    ::. . .   . :     ::.  .  :
Qy    156 CYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 215

Db    673 flnvleaiepgvvcaghdnnqpdsfaallsslnelgerqlvhvvkwakalpgfrnlhvdd 732
           |.:|||  | |:  :  :  | . |... ||  |.:::|||:: ||| :||| :| : |
Qy    216 VLTLLEAEPPHVL-ISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 273

Db    733 qmaviqyswmglmvfamgwrsftnvnsrmlyfapdlvfn-eyrmhksrmysqcvrmrhls 791
          |: :::  ||  ::  .: |||  .  .::: ||||||.: :       :      :   :
Qy    274 QVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDMLLATT 331

Db    792 qefgwlqitpqeflcmkalllfs 814
            |  | :   |:||:||::|::
Qy    332 SRFRELKLQHKEYLCVKAMILLN 354


Search completed: Thu Dec 18 08:53:55 1997
Job time : 83 secs.
```

```
***********************************************************************

                          MPSRCH                          (TM)

***********************************************************************

         Release 2.1D John F. Collins, Biocomputing Research Unit.
         Copyright (c) 1993, 1994, 1995  University of Edinburgh, U K
                  Distribution rights by IntelliGenetics, Inc.

MPsrch_pp   protein - protein database search, using Smith-Waterman algorithm

Run on:            Thu Dec 18 08:41:12 1997,  MasPar time 11.31 Seconds
                                              518.308 Million cell updates/sec
Tabular output not generated.

Title:             >US-08-826-361A-5
Description:       (1-477) from US08826361A.pep
Perfect Score:     3532
Sequence:          1 MNYSIPSNVTNLEGGPGRQT.........ECSPAEDSKSKEGSQNPQSQ 477

Scoring table:     PAM 150
                   Gap 11

Searched:          101610 seqs, 12294212 residues

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database:          a-geneseq28
                     1:part1 2:part2 3.part3 4.part4 5.part5 6:part6 7:part7
                     8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
                     14:part14 15:part15 16:part16 17:part17 18:part18
                     19:part19 20:part20 21.part21

Statistics:        Mean 35.707;  Variance 150.730;  scale 0.237

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES
                 %
Result         Query
  No.  Score   Match Length DB  ID              Description           Pred. No.
-------------------------------------------------------------------------------
    1   3520    99.7    485 21  W14724          Human oestrogen recep 0.00e+00
    2   3212    90.9    495 21  W14723          Rat oestrogen recepto 0.00e+00
    3   3191    90.3    495 21  W14725          Mouse oestrogen recep 0.00e+00
    4   1671    47.3    595  2  P70543          Human oestrogen recep 2.32e-152
    5   1424    40.3    334  2  P80925          Sequence of the human 1.82e-127
    6    807    22.8    433  2  P80931          Sequence of human est 8.04e-66
    7    800    22.7    521  2  P80930          Sequence of human est 3.96e-65
    8    614    17.4    777  2  P80919          Sequence of the prima 7.62e-47
    9    576    16.3    902  3  R12224          Rat androgen receptor 3.85e-43
   10    570    16.1    902  1  P93110          Rat androgen receptor 1.48e-42
   11    570    16.1    902  1  P91006          Rat androgen receptor 1.48e-42
   12    570    16.1    919  3  R12223          Human androgen recept 1.48e-42
   13    567    16.1    919  1  P90996          Human androgen recept 2.89e-42
   14    567    16.1    919 21  W14783          Androgen receptor     2.89e-42
   15    566    16.0    919  1  P93109          Human androgen recept 3.62e-42
   16    563    15.9     77 10  R53289          Human oestrogen recep 7.08e-42
   17    546    15.5     72  1  P91994          Part of oestrogen rec 3.16e-40
   18    545    15.4    630  3  R12230          TrpE/androgen recepto 3.96e-40
   19    536    15.2     68 13  R71566          Partial human estroge 2.95e-39
   20    533    15.1    533 13  R72483          Human H-2RIIBP.       5.76e-39
   21    533    15.1    533  8  R39468          hRXR-beta1.           5.76e-39
   22    529    15.0    446  2  R11937          H2RIIBF nuclear hormo 1.40e-38
   23    529    15.0    446 13  R72484          H-2RIIBP.             1.40e-38
   24    525    14.9    448  8  R39467          mRXR-beta.            3.43e-38
   25    528    14.9    510  8  R39469          hRXR-beta2.           1.76e-38
   26    521    14.8    984  2  P80927          Sequence of the human 8.35e-38
   27    514    14.6    467  8  R39470          mRXR-alpha.           3.97e-37
   28    514    14.6    467  3  R13799          Mouse RXR-alpha retin 3.97e-37
   29    509    14.4    462  3  R13798          Human RXR-alpha retin 1.21e-36
   30    489    13.8    463  8  R39471          mRXR-gamma.           1.02e-34
   31    486    13.8    463  3  R13800          Mouse RXR-gamma retin 1.99e-34
   32    473    13.4    463 20  W00605          Human retinoid X rece 3.55e-33
   33    471    13.3    598  9  R48631          Sequence of nuclear r 5.52e-33
   34    449    12.7    447  4  R20466          Mouse retinoic acid r 7.11e-31
   35    450    12.7    458  2  R10549          Skin-specific murine  5.70e-31
   36    449    12.7    458  4  R20465          Mouse retinoic acid r 7.11e-31
   37    444    12.6    448  2  R10548          Murine Retinoic Acid  2.14e-30
   38    444    12.6    448 16  R84725          Murine retinoic acid  2.14e-30
   39    442    12.5    448 10  R55128          Retinoic acid recepto 3.32e-30
   40    441    12.5    448 17  R85315          Human hepatoma retino 4.14e-30
   41    441    12.5    448 12  R66764          Human RAR-beta        4.14e-30
   42    441    12.5    454  2  R10182          Recombinant human gam 4.14e-30
   43    441    12.5    454  2  R10550          Human Retinoic Acid R 4.14e-30
   44    441    12.5    454 16  R84723          Human retinoic acid r 4.14e-30
   45    443    12.5    458 16  R84726          Murine retinoic acid  2.67e-30

                              ALIGNMENTS

RESULT    1
ID   W14724 standard; Protein; 485 AA.
AC   W14724;
DT   08-JUN-1997  (first entry)
DE   Human oestrogen receptor beta (ER-beta).
KW   Orphan receptor; oestrogen receptor beta; ER-beta;
KW   nuclear receptor; prostate cancer; benign prostatic hyperplasia;
KW   ovary cancer; cardiovascular disease; osteoporosis;
KW   environment; pollutant.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   Domain          104..169
FT   /label= DNA-binding_domain
FT   Domain          260..457
FT   /label= Ligand-binding_domain
PN   WO9709348-A2.
PD   13-MAR-1997.
PF   09-SEP-1996; E03933.
PR   08-SEP-1995; GB-018272.
PR   15-MAR-1996; GB-005550.
PR   11-APR-1996; GB-007532.
PR   08-MAY-1996; GB-009576.
PA   (KARO-) KARO BIO AB.
PI   Enmark E,  Gustafsson J,  Kuiper GG;
DR   WPI; 97-192842/17.
DR   N-PSDB; T62843.
PT   New isolated oestrogen receptor beta - used to develop prods. for
PT   treating e.g. cancers, CNS diseases, osteoporosis or cardiovascular
PT   disease
PS   Claim 1; Fig 13A; 45pp; English.
CC   A novel human orphan receptor (W14724) is related to the known
CC   oestrogen receptor ER-alpha, and has been designated ER-beta.
CC   It is an oestrogen receptor-related nuclear receptor.  The ER-beta
CC   amino acid sequence was deduced from a cDNA clone (T62843) isolated
CC   from a human testis cDNA library.  Rat, human and mouse ER-beta
CC   (W14723-25) can be used to isolate molecules for use in the
CC   treatment of cardiovascular diseases, central nervous system
CC   diseases, osteoporosis, prostate or ovarian cancer or benign
CC   prostatic hyperplasia and to test environmental chemicals for
CC   oestrogenic activity.
SQ   Sequence   485 AA;

  Query Match             99.7%;  Score 3520;  DB 21;  Length 485;
  Best Local Similarity   99.8%;  Pred. No. 0.00e+00;
  Matches   476;  Conservative    0;  Mismatches   1;  Indels   0;  Gaps   0;

Db        9 mnysipsnvtnleggpgrqttspnvlwptpghlsplvvhrqlshlyaepqkspwcearsl 68
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        1 MNYSIPSNVTNLEGGPGRQTTSPNVLWPTPGHLSPLVVHRQLSHLYAEPQKSPWCEARSL 60

Db       69 ehtlpvnretlkrkvsgnrcaspvtgpgskrdahfcavcsdyasgyhygvwscegckaff 128
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       61 EHTLPVNRETLKRKVSGNPCASPVTGPGSKRDAHFCAVCSDYASGYHYGVWSCEGCKAFF 120

Db      129 krsiqghndyicpatnqctidknrrkscqacrlrkcyevgmvkcgsrrercgyrlvrrqr 188
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      121 KRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQR 180

Db      189 sadeqlhcagkakrsgghaprvrellldalspeqlvltlleaepphvlisrpsapfteas 248
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      181 SADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQLVLTLLEAEPPHVLISRPSAPFTEAS 240

Db      249 mmmsltkladkelvhmiswakkipgfvelslfdqvrllescwmevlmmglmwrsidhpgk 308
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      241 MMMSLTKLADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGK 300

Db      309 lifapdlvldrdegkcvegileifdmllattsrfrelklqhkeylcvkamillnssmypl 368
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      301 LIFAPDLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPL 360

Db      369 vtatqdadssrklahllnavtdalvwviaksgissqqqsmrlanllmllshvrhasnkgm 428
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      361 VTATQDADSSRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGM 420

Db      429 ehllnmkcknvvpvydlllemlnahvlrgckssitgsecspaedskskegsqnlqsq 485
            ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Qy      421 EHLLNMKCKNVVPVYDLLLEMLNAHVLRGCKSSITGSECSPAEDSKSKEGSQNPQSQ 477

RESULT    2
ID   W14723 standard; Protein; 485 AA.
AC   W14723;
DT   08-JUN-1997  (first entry)
DE   Rat oestrogen receptor beta (ER-beta).
KW   Orphan receptor; oestrogen receptor beta; ER-beta;
KW   nuclear receptor; prostate cancer; benign prostatic hyperplasia;
KW   ovary cancer; cardiovascular disease; osteoporosis;
KW   environment; pollutant.
OS   Rattus sp.
FH   Key             Location/Qualifiers
FT   Domain          104..169
FT   /label= DNA binding_domain
FT   Domain          259..457
FT   /label= Ligand-binding_domain
PN   WO9709348-A2.
PD   13-MAR-1997.
PF   09-SEP-1996; E03933.
PR   08-SEP-1995; GB-018272
PR   15-MAR-1996; GB-005550.
PR   11-APR-1996; GB-007532.
PR   08-MAY-1996; GB-009576.
PA   (KARO-) KARO BIO AB.
PI   Enmark E,  Gustafsson J,  Kuiper GG;
DR   WPI; 97-192842/17.
DR   N-PSDB; T62842.
PT   New isolated oestrogen receptor beta - used to develop prods. for
PT   treating e.g. cancers, CNS diseases, osteoporosis or cardiovascular
PT   disease
PS   Claim 1; Fig 1; 45pp; English.
CC   A novel rat orphan receptor (W14723) is related to the known
CC   oestrogen receptor ER-alpha, and has been designated ER-beta.
CC   It is an oestrogen receptor-related nuclear receptor.  The ER-beta
CC   amino acid sequence was deduced from a cDNA clone (T62842) isolated
CC   from a rat prostate cDNA library.  Rat, human and mouse ER-beta
CC   (W14723-25) can be used to isolate molecules for use in the
CC   treatment of cardiovascular diseases, central nervous system
CC   diseases, osteoporosis, prostate or ovarian cancer or benign
CC   prostatic hyperplasia and to test environmental chemicals for
CC   oestrogenic activity.
SQ   Sequence   485 AA;

  Query Match             90 9%;  Score 3212,  DB 21,  Length 485;
  Best Local Similarity   88 5%;  Pred  No  0 00e+00;
  Matches   422;  Conservative   35;  Mismatches   20;  Indels   0;  Gaps   0;

Db        9 mnysvpgstsnldggpvilstspnvlwptsghlsplathcqssllyaepqkspwcearsl 68
            ||||:|:: :||:||| | :|||||||||:||||||   | | | ||||||||||||||||
Qy        1 MNYSIPSNVTNLEGGPGRQTTSPNVLWPTPGHLSPLVVHRQLSHLYAEPQKSPWCEARSL 60

Db       69 ehtlpvnretlkrklsgsscaspvtspnakrdahfcpvcsdyasgyhygvwscegckaff 128
            ||||||||||||||.||. ||||||.| :||||||||.||||||||||||||||||||||
Qy       61 EHTLPVNRETLKRKVSGNPCASPVTGPGSKRDAHFCAVCSDYASGYHYGVWSCEGCKAFF 120

Db      129 krsiqghndyicpatnqctidknrrkscqacrlrkcyevgmvkcgsrrercgyrivrrqr 188
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Qy      121 KRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQR 180

Db      189 ssseqvhclskakrngghaprvkelllstlspeqlvltlleaeppnvlvsrpsmpfteas 248
            |: ||:||  ||||:|||||||:||||. ||||||||||||||||:||:|||| ||||||
Qy      181 SADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQLVLTLLEAEPPHVLISRPSAPFTEAS 240

Db      249 mmmsltkladkelvhmigwakkipgfvelslldqvrllescwmevlmvglmwrsidhpgk 308
            |||||||||||||||||:|||||||||||||:||||||||||||||||:|||||||||||
Qy      241 MMMSLTKLADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGK 300

Db      309 lifapdlvldrdegkcvegileifdmllattsrfrelklqhkeylcvkamillnssmypl 368
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      301 LIFAPDLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPL 360

Db      369 asanqeaessrklthllnavtdalvwviaksgissqqqsvrlanllmllshvrhisnkgm 428
             :| |:|:|||||:|||||||||||||||||||||||||:||||||||||||| ||||||
Qy      361 VTATQDADSSRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGM 420

Db      429 ehllsmkcknvvpvydlllemlnahtlrgykssisgsecsstedsknkessqnlqsq 485
            ||||:|||||||||||||||||||| ||| ||||:|||||::||||:||:||| |||
Qy      421 EHLLNMKCKNVVPVYDLLLEMLNAHVLRGCKSSITGSECSPAEDSKSKEGSQNPQSQ 477

RESULT    3
ID   W14725 standard; Protein; 485 AA.
AC   W14725;
DT   08-JUN-1997  (first entry)
DE   Mouse oestrogen receptor beta (ER-beta)
KW   Orphan receptor; oestrogen receptor beta; ER-beta;
KW   nuclear receptor; prostate cancer; benign prostatic hyperplasia;
KW   ovary cancer; cardiovascular disease; osteoporosis;
KW   environment; pollutant.
OS   Mus sp.
FH   Key             Location/Qualifiers
FT   Domain          104  169
FT   /label= DNA-binding_domain
FT   Domain          260..457
FT   /label= Ligand-binding_domain
PN   WO9709348-A2.
PD   13-MAR-1997
PF   09-SEP-1996; E03933.
PR   08-SEP-1995; GB-018272.
PR   15-MAR-1996; GB-005550.
PP   11-APR-1996; GB-007532.
PF   08-MAY-1996, GB-009576
PA   (KARO-) KARO BIO AB
PI   Enmark E,  Gustafsson J,  Kuiper GG;
DR   WPI; 97-192842/17.
DR   N-PSDB; T62844.
PT   New isolated oestrogen receptor beta - used to develop prods. for
PT   treating e.g. cancers, CNS diseases, osteoporosis or cardiovascular
PT   disease
PS   Claim 1, Fig 14A, 45pp, English.
CC   A novel mouse orphan receptor (W14725) is related to the known
CC   oestrogen receptor ER-alpha, and has been designated ER-beta.
CC   It is an oestrogen receptor-related nuclear receptor. The ER-beta
CC   amino acid sequence was deduced from a cDNA clone (T62844).
CC   Rat, human and mouse ER-beta (W14723-25) can be used to isolate
CC   molecules for use in the treatment of cardiovascular diseases,
CC   central nervous system diseases, osteoporosis, prostate or ovarian
CC   cancer or benign prostatic hyperplasia and to test environmental
CC   chemicals for oestrogenic activity.
SQ   Sequence   485 AA;

  Query Match             90.3%;  Score 3191;  DB 21;  Length 485;
  Best Local Similarity   88.1%;  Pred. No. 0.00e+00;
  Matches  420;  Conservative   35;  Mismatches   22;  Indels    0;  Gaps    0;

Db          9 mnysvpsstgnleggpvrqtaspnvlwptsghlsplathgqssllyaepqkspwcearsl 68
              ||||:||:  |||||| |||:||||||||:||||||  | | | |||||||||||||||
Qy          1 MNYSIPSNVTNLEGGPGRQTTSPNVLWPTPGHLSPLVVHRQLSHLYAEPQKSPWCEARSL 60

Db         69 ehtlpvnretlkrklggsgcaspvtspstkrdahfcavcsdyasgyhygvwscegckaff 128
              ||||||||||||||::|: ||||||:|:||||||||||||||||||||||||||||||||
Qy         61 EHTLPVNRETLKRKVSGNRCASPVTGPGSKRDAHFCAVCSDYASGYHYGVWSCEGCKAFF 120

Db        129 krsiqghndyicpatnqctidknrrksgqacrlrkcyevgmvkcgsrrercgyrivrrqr 188
              |||||||||||||||||||||||||||:||||||||||||||||||||||||||||:||||
Qy        121 KRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQR 180

Db        189 saseqvhclnkakrtsghtprvkelllnslspeqlvltlleaeppnvlvsrpsmpfteas 248
              || ||:||  ||||::||:||·||||··||||||||||||||||·||:|||| ||||||
Qy        181 SADEQLHCAGKAKRSGGHAPPVRELLLDALSPEQLVLTLLEAEPPHVLISRPSAPFTEAS 240

Db        249 mmmsltkladkelvhmigwakkipgfvelslldqvrllescwmevlmvglmwrsidhpgk 308
              |||||||||||||||||:||||||||||||||:|||||||||||||||:||||||||||||
Qy        241 MMMSLTKLADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGK 300

Db        309 lifapdlvldrdegkcvegileifdmllattarfrelklqhkeylcvkamillnssmyhl 368
              |||||||||||||||||||||||||||||||:||||||||||||||||||||||||||| |
Qy        301 LIFAPDLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPL 360

Db        369 atasqeaessrklthllnavtdalvwvisksrissqqqsvrlanllmllshvrhisnkgm 428
               ||:|:|:||||:||||||||||||||||:|| |||||||:|||||||||||||| |||||
Qy        361 VTATQDADSSRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGM 420

Db        429 ehllsmkcknvvpvydlllemlnahtlrgykssisgsgccstedskskegsqnlqsq 485
              ||||:||||||||||||||||||||| ||| ||||:|| | ::|||||||||||| |||
Qy        421 EHLLNMKCKNVVPVYDLLLEMLNAHVLRGCKSSITGSECSPAEDSKSKEGSQNPQSQ 477


RESULT    4
ID   P70543 standard; protein; 595 AA
AC   P70543;
DT   24-FEB-1991  (first entry)
DE   Human oestrogen receptor protein.
KW   Human oestrogen receptor protein; expression system; assay.
OS   Homo sapiens.
PN   WO8705049-A
PD   27-AUG-1987.
PF   18-FEB-1987; U00341.
PR   20-FEB-1986; US-833829.
PA   (CALB-) California Biotechnology Inc.
PI   Shine J;
DR   WPI; 87-250213/35.
DR   N-PSDB; N70880.
PT   Expression system for vertebrate steroid receptor protein - comprising
PT   DNA sequence encoding the protein linked to control sequences in
PT   eucaryotic hosts
PS   Disclosure; Fig 1-1 - 1-2; 26pp; English
CC   steroid receptor protein.
CC   The protein is expressed in an expression system from cDNA operably
CC   linked to control sequences compatible with eukaryotic host cells.
CC   This method allows expression under conditions which favour appropriate
CC   post-translational processing.  It also produces large amts. of purified
CC   protein useful in the design of agonist and antagonist cpds. for study
CC   of the mechanism of action of the steroid binding proteins in general.
CC   and for use in diagnostic assays for the proteins or antibodies to
```

```
CC   them. These assays are important in, eg the diagnosis of tumour
CC   sensitivities to steroid metabolism. Suitable host cells are VERO,
CC   HeLa and CHO cells.
SQ   Sequence   595 AA;

  Query Match             47.3%;  Score 1671;  DB 2;  Length 595;
  Best Local Similarity   56.3%;  Pred. No. 2.32e-152;
  Matches  227;  Conservative   98;  Mismatches   61;  Indels   17;  Gaps   13;

Db        178 saketrycavcndyasgyhygvwscegckaffkrsiqghndymcpatnqctidknrrksc 237
              | ..::|||:|||||||||||||||||||||||||||||||||.|||||||||||||||
Qy         89 SKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSC 148

Db        238 qacrlrkcyevgmmkggirkdrrggrmlkhkrqrddgegrgevgsagdmraanlwpsplm 297
              |||||||||||||:| | |::| | |:::  |||   | :   :: : :|::     :|
Qy        149 QACRLRKCYEVGMVKCGSRRERCGYRLVR--RQRSADE-QLHCAGKA-KRSGG--HAP-P 201

Db        298 ikrskknslalsltadqmvsalldaepp-ilyseydptrpfseasmmglltnladrelvh 356
              : |      |  .|...|:| .||.|||| .| |    |: ||:|||||  ||:|||:||||
Qy        202 V-RELL--LD-ALSPEQLVLTLLEAEPPHVLISR--PSAPFTEASMMMSLTKLADKELVH 255

Db        357 minwakrvpgfvdltlhdqvhllecawleilmiglvwrsmehpgkllfapnllldrnqgk 416
              ||:|||:.||||.|.| |||.|||  |.|.||.||.|||..|| ||.|||:|:|||::||
Qy        256 MISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGK 315

Db        417 cvegmveifdmllatssrfrmmnlqgeefvclksiillnsgvytflsstlksleekdhih 476
              ||||::||||||||'||.||||  ..||  |..|.|..||||||..| ...:|    .   .
Qy        316 CVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSRKLA 374

Db        477 rvldkitdtlihlmakagltlqqqhqrlaqlllilshirhmsnkgmehlysmkcknvvpl 536
              . |   :||  |     ||. .: |||  ||| ||.:|||:|| |||||||| :|||||||:
Qy        375 HLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPV 434

Db        537 ydlllemldahrlha-ptsrggasveetdqshlatagstsshs 578
              |||:||||:|| |::  :|  |:     :::|  :  || :::|
Qy        435 YDLLLEMLNAHVLRGCKSSITGSECSPAEDSK-SKEGSQNPQS 476


RESULT    5
ID   P80925 standard, protein, 334 AA.
AC   P80925;
DT   29-DEC-1990  (first entry)
DE   Sequence of the human oestrogen receptor
KW   Hormone receptor; hormone-binding; transcription activation;.
OS   Homo sapiens.
PN   WO8803168-A.
PD   05-MAY-1988.
PF   23-OCT-1987; U02782.
PR   20-OCT-1987; US-108471.
PA   (SALK) Salk Inst for Biol Stud.
PI   Evans RM, Weinberger CA, Hollenberg SM, Giguere V;
DR   WPI; 88-133242/19.
PT   Recombinant DNA encoding hormone recpetors -
PT   comprising glucocorticoid, mineralocorticoid, thyroid hormone
PT   and novel hormone receptors
PS   Example; Fig III-2; 243pp; English.
CC   DNAs encoding hormone receptors and the hormone receptors themselves are
CC   claimed. The DNA can be used to make the hormone receptor proteins and
CC   functional modified forms in quantities not previously possible. The
CC   receptor proteins can be used to screen cpds  for receptor-agonist or
CC   receptor-antagonist activity. They can also be used in diagnostic assays.
SQ   Sequence   334 AA;

  Query Match             40.3%,  Score 1424,  DB 2,  Length 334;
  Best Local Similarity   57.4%,  Pred. No. 1.52e-127;
  Matches  193,  Conservative   60,  Mismatches   47,  Indels   16;  Gaps   12;

Db          1 ycavcndyasgyhygvwscegckaffkrs-qghndymcpatnqctidknrrkscqacrlr 59
              :||||:|||||||||||||||||||||||| ||||||:|||||||||||||||||||||||
Qy         95 FCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRLR 154

Db         60 kcyrvgmmkggirkdrrggrmlkhkrqrddgegrgevgsagdmraanlwpsplmikrskk 119
              ||| |||.| | |.:| | |...  |||    | .    :: : :|::     :|  : |
Qy      155 KCYEVGMVKCGSRRERCGYRLVR--RQRSADE-QLHCAGKA-KRSGG--HAP-RV-RELL 206

Db      120 sslalsltadqmvsalldaeppilyseydptrpfseasmmglltnladrelvhminwak 178
               |   :|:::|:|  :||:||||  :|  |    |: ||:|||||   ||:|||:||||||:|||
Qy      207 --LD-ALSPEQLVLTLLEAEPPHVLISR--PSAPFTEASMMMSLTKLADKELVHMISWAK 261

Db      179 rvpgfvdltlhdqvhllecawleilmiglvwrsmehpvkllfapnllldrnqgkcvegmv 238
            ::||||:|:|  |||:|||   |:|:||:||:|||   || || ||| | |||  ||||||
Qy      262 KIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGIL 321

Db      239 eifdmllatssrfrmmnlqgeefvclksiillnsgvytflsstlksleekdhihrvldki 298
            ||||||||:||||  ::||   |..|.|..|||||..|  ....|   .  .      . ..|. .
Qy      322 EIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSRKLAHLLNAV 380

Db      299 tdtlihlmakaqltlqqqhqrlaqlllilshirhms 334
            ||:|: ::||:|:: |||  ||| ||::|||:|| |
Qy      381 TDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHAS 416


RESULT    6
ID   P80931 standard; protein; 433 AA
AC   P80931;
DT   29 DEC 1990  (first entry)
DE   Sequence of human estrogen-related receptor protein (hERR2)
KW   Hormone receptor; hormone-binding; transcription activation.
OS   Homo sapiens.
PN   WO8803168-A
PD   05-MAY-1988
PF   23-OCT-1987; U02782.
PR   20-OCT-1987; US-108471
PA   (SALK) Salk Inst for Biol Stud
PI   Evans RM, Weinberger CA, Hollenberg SM, Giguere V;
DR   WPI; 88-133242/19
DR   N-PSDB; N80922.
PT   Recombinant DNA encoding hormone recpetors -
PT   comprising glucocorticoid, mineralocorticoid, thyroid hormone
PT   and novel hormone receptors
PS   Claim 17; Fig V-2(B)-1 and  2; 243pp; English.
CC   DNAs encoding hormone receptors and the hormone receptors themselves are
CC   claimed. The DNA can be used to make the hormone receptor proteins and
CC   functional modified forms in quantities not previously possible. The
CC   receptor proteins can be used to screen cpds  for receptor-agonist or
CC   receptor-antagonist activity  They can also be used in diagnostic assays
CC   Also claimed is pure DNA encoding oestrogen-related receptors hERR1 and
CC   hERR2 and hERR1 and hERR2  The new hERR1 and hERR2 receptors will provide
CC   the basis for development of an assay system that will lead to the
CC   identification of novel hormones
SQ   Sequence   433 AA;

  Query Match              22.8%,  Score 807,  DB 2,  Length 433;
  Best Local Similarity  39.5%;  Pred  No  8.04e-66;
  Matches   140,  Conservative   84,  Mismatches 108;  Indels  22,  Gaps  14,

Db      101 rlclvcgdiasgyhygvasceackaffkrtiqgnieyscpatneceitkrrrkscqacrf 160
            ::| ||:| ||||||||| |||:||||||||:|||: :| |||||:| | | |||||||||:
Qy       94 HFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRL 153

Db      161 mkclkvgmlkegvrldrv--rggrqky-krrldsenspylslqisppakkpltkivs-- 214
             ||  |||:| | | :|   |   ||:      :|          :|    :|
Qy      154 RKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPE 213

Db      215 yl-l-taepdkiyam-ppddvpedikalttlcdladrelvfliswakhipgfsnltlgd 271
             | | : |::     : :     |   :      : :|  |||:|||  :|||||  ||||  :|:| |
Qy      214 QLVLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 273

Db      272 qmsllqsawmeililgivyrslpyddklayaedyimdeehsrlv-glleiyrailqlvrr 330
            |: ||:| |||:|  |  ||        || :| |  :  |       | |.|| .  .|   |
Qy      274 QVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSR 333

Db      331 ykklkvekeefvmlkaialansdsmy--i--enleavqklqdllh--ealq-dyels- 380
            :: ||::  |:: :||: | || |||   :      .. .. .||  ||.    .||     |
Qy      334 RRELKLQHKEYLCVKAMILLNS SMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKSG 392
```

```
Db      381 -qrheeprragkllltlpllrqtaakavqhfysvklqgkvpmhklflemleakv 433
               .::: | :.||: |: .|... |...|. ..|      ||:   |:||||:| |
Qy      393 ISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNAHV 440


RESULT    7
ID   P80930 standard; protein; 521 AA.
AC   P80930,
DT   10-MAR-1993  (revised)
DT   29-DEC-1990  (first entry)
DE   Sequence of human estrogen-related receptor protein (hERR1)
KW   Hormone receptor, hormone-binding, transcription activation.
OS   Homo sapiens.
PN   WO8803168-A.
PD   05-MAY-1988.
PF   23-OCT-1987; U02782.
PR   20-OCT-1987; US-108471.
PA   (SALK) Salk Inst for Biol Stud.
PI   Evans RM, Weinberger CA, Hollenberg SM, Giguere V;
DR   WPI; 88-133242/19.
DR   N-PSDB; N80921
PT   Recombinant DNA encoding hormone recpetors -
PT   comprising glucocorticoid, mineralocorticoid, thyroid hormone
PT   and novel hormone receptors
PS   Claim 17, Fig V-1(B)-1 and -2; 243pp, English.
CC   DNAs encoding hormone receptors and the hormone receptors themselves are
CC   claimed  The DNA can be used to make the hormone receptor proteins and
CC   functional modified forms in quantities not previously possible. The
CC   receptor proteins can be used to screen cpds. for receptor-agonist or
CC   receptor-antagonist activity  They can also be used in diagnostic assays.
CC   Also claimed is pure DNA encoding oestrogen-related receptors hERR1 and
CC   hERR2 and hERR1 and hERR2  The new hERR1 and hERR2 receptors will provide
CC   the basis for development of an assay system that will lead to the
CC   identification of novel hormones.
SQ   Sequence   521 AA;

  Query Match              22.7%;  Score 800;  DB 2;  Length 521;
  Best Local Similarity  39.9%;  Pred. No. 3.96e-65;
  Matches   141;  Conservative   74;  Mismatches 128;  Indels  10;  Gaps   7;

Db      174 rlclvcgdvasgyhygvasceackaffkrtiqgsieyscpasneceitkrrrkacqacrf 233
            ::| ||:| |||||||| |||:|||||||||:|||   :| |||:|:| | | |||:||||||:
Qy       94 HFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRL 153

Db      234 tkclrvgmlkegvrldrvrggrqkykrrpevdplpfpg-p-fpagplavaggprktaapv 291
             ||  |||:| | | :| ||    ..:.  . |   :| :  ::|  :           | :
Qy      154 RKCYEVGMVKCGSRRERC-GYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSP 212

Db      292 nalvshllvvepeklyampdpagpdghlpavatlcdlfdreivvtiswaksipgfsslsl 351
            . ||  ||  || :  :. |::|      . . .|  | |:|:|  |||||  ||||   |||
Qy      213 EQLVLTLLEAEPPHVL-ISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSL 271

Db      352 sdqmsvlqsvwmevlvlgvaqrslplqdelafaedlvld-eegaraaglgelgaallqlv 410
             ||    | : ||:||  |.  ||.       | || |||||  :||      |: |:    ||
Qy      272 FDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATT 331

Db      411 rrlqalrlereeyvllkalalansdsvhiedeprlwsscekllh--eal---leyeaqra 465
             :..:|.:.. ||. .:||. | ||        .     :    | ||| |  .:|.    |      :::
Qy      332 SRRRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKS 391

Db      466 gpgggaerrragrllltlpllrqtagkvlahfygvklegkvpmhklflemlea 518
             |   :    : :  ||. |. :|::: | : :       :      ||     : :||:|  |
Qy      392 GISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNA 444


RESULT    8
ID   P80919 standard, protein, 777 AA.
AC   P80919;
DT   29-DEC-1990  (first entry)
DE   Sequence of the primary protein sequence of human glucocorticoid
DE   receptor (hGR)
KW   Hormone receptor, hormone-binding, transcription activation.
OS   Homo sapiens.
```

```
PN   WO8803168-A.
PD   05-MAY-1988.
PF   23-OCT-1987; U02782.
PR   20-OCT-1987; US-108471.
PA   (SALK) Salk Inst for Biol Stud.
PI   Evans RM, Weinberger CA, Hollenberg SM, Giguere V;
DR   WPI; 88-133242/19.
DR   N-PSDB; N80916.
PT   Recombinant DNA encoding hormone recpetors -
PT   comprising glucocorticoid, mineralocorticoid, thyroid hormone
PT   and novel hormone receptors
PS   Claim 24; Fig I-2(1) to (2); 243pp; English
CC   DNAs encoding hormone receptors and the hormone receptors themselves are
CC   claimed. The DNA can be used to make the hormone receptor proteins and
CC   functional modified forms in quantities not previously possible. The
CC   receptor proteins can be used to screen cpds. for receptor-agonist or
CC   receptor-antagonist activity. They can also be used in diagnostic assays.
SQ   Sequence   777 AA;

  Query Match            17.4%;  Score 614;  DB 2,  Length 777,
  Best Local Similarity  33.4%;  Pred. No. 7.62e-47;
  Matches   97;  Conservative   84;  Mismatches  93;  Indels  16,  Gaps  15;

Db     397 smrpdvssppssstattgp-p-pklclvcsdeasgchygvltcgsckvffkravegqhn 454
            :::  ||:    :|: :  |:    : :| |||| ||| |||| :| :|| ||||:::|:::
Qy      70 ILKRKVSGNRCASPVTGPGSKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHND 129

Db     455 ylcagrndciidkirrkncpacryrkclqagmnl-carktk-k-kikgiqqattg-v-sq 509
           |:|:: |:| ||| |||:| ||| ||| : ||       :|. :     ..    |...    :
Qy     130 YICPATNQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCA 189

Db     510 etsenpgnkt-ivpatlpq-ltpt-lv-slleviepevlyagydssvpdstwrimttlnm 565
              :   :|  :   |     | : |:|  || :|||     | || :    ::   :::   :| :| :
Qy     190 GKAKRSGGHAPRVRELLLDALSPEQLVLTLLEAEPPHVLISRPSAPFTEAS--MMMSLTK 247

Db     566 lqgrqviaavkwakaipgfrnlhlddqmtllqyswmflmafalgwrsyrqssanllcfap 625
           |: ::::   : ||| |||| :| | ||: ||:   || ::    :| |||   ::: :|: |||
Qy     248 LADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPG-KLI-FAP 305

Db     626 dliin-eqrmtlpcmydqckhmlyvsselhrlqvsyeeylcmktllllss 674
           |':::  ::  .   .   .    .|   .| :: |  .    ||||:|:::||:|
Qy     306 DLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNS 355

RESULT    9
ID   R12224 standard; Protein; 902 AA.
AC   R12224;
DT   20-AUG-1991  (first entry)
DE   Rat androgen receptor.
KW   rAR; DNA-binding protein, steroid hormone.
OS   Rattus rattus.
FH   Key             Location/Qualifiers
FT   Domain          540..610
FT   /label= DNA-binding domain
FT   /note= "cysteine-rich"
PN   WO9107423-A.
PD   30-MAY-1991.
PF   19-OCT-1990; U06015.
PR   17-NOV-1989; US-438775.
PA   (ARCH-) ARCH DEV CORP.
PI   Liao S, Chang C;
DR   WPI; 91-178048/24.
DR   N-PSDB; Q12002.
PT   Androgen receptor and TR2 DNA binding proteins - DNA sequences
PT   and antibodies for detection and quantification methods
PS   Claim 25; Fig 3; 79pp; English
CC   This sequence was deduced from a cDNA clone isolated by screening
CC   a rat ventral prostate lambda gt11 library in E.coli Y1090. The
CC   sequence is very similar to that of human AR and in
CC   the DNA-binding domain it is identical to that of hAR DNA-binding
CC   domain. Homology comparisons with other known steroid receptors
CC   indicate that rAR is more closely related to glucocorticoid,
CC   mineralo-corticoid and progesterone receptors than to v-erb-A or to
CC   receptors for oestrogen, vitamin D and thyroid hormones.
SQ   Sequence   902 AA;

  Query Match            16.3%;  Score 576;  DB 3;  Length 902;
  Best Local Similarity  30.1%;  Pred. No. 3.85e-43;
  Matches  108;  Conservative  103;  Mismatches 126;  Indels  22;  Gaps  19;

Db     542 clicgdeasgchygaltcgsckvffkraaegkqkylcasrndctidkfrrkncpscrlrk 601
            | :|:| ||| |||   :| :|| ||||: :|    |:|:: |:||||| |||:| :|||||
Qy      96 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 155

Db     602 cyeagmtl-garklk---klgnlklqeegenssagsptedpsqkmtvshi--egyecqpi 655
            ||| ||    |:|: :     :|          : :   || :     ..   |  .  ..   : :
Qy     156 CYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 215

Db     656 flnvleaiepgvvcaghdnnqpdsfaallsslnelgerqlvhvvkwakalpgfrnlhvdd 715
             | :||!  | |.             |  . |:  . ||  |: : |||:. ||| .||| .| : |
Qy     216 VLTLLEAEPPHVL-ISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 273

Db     716 qmavieyswmglmvfamgwrsftnvnsrmlyfapdlvfn-eyrmhksrmysqcvrmrhls 774
           |: ::|  ||  ...     |||.  .    .  ||||||  .  .  :      :      :   :
Qy     274 QVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDMLLATT 331

Db     775 qefgwlqitpqeflcmkal-ll-fsiip-vdglknqkffdel-rm-nyike-ldriiack 828
             |   | :    |:||:||: ||   |: | | | :   :         | :: |  :  : | ::'|
Qy     332 SRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKS 391

Db     829 rknptscsrrfyqltklldsv-qpiarelqqftfdllikshmvsvdfpemmaeiisvqv 886
              ::   | |:  ' .||  | :.   .  .::: ...    | ::|:| : ::: |::: :|
Qy     392 GISSQQQSMRLANLLMLLSHVRHASNKGMEHL-LNMKCK-NVVPV-Y-DLLLEMLNAHV 446

RESULT    10
ID   P93110 standard; protein; 902 AA.
AC   P93110;
DT   19-MAR-1990  (first entry)
DE   Rat androgen receptor.
KW   Rat androgen receptor, monoclonal antibody, ployclonal antibody; cancer.
OS   Rattus rattus.
PN   WO8909791-A.
PD   19-OCT-1989.
PF   13-APR-1989; U01548.
PR   14-APR-1988; US-182646.
PA   (UYNC-) University of North Carolina.
PI   French FS, Wilson EM, Joseph DR, Lubahn DB;
DR   WPI; 89-324206/44.
DR   N-PSDB; N91773.
PT   DNA encoding androgen receptor protein - useful for transforming
PT   eukaryotic hosts for protein expression and subsequent antibody prodn.
PS   Disclosure; Fig. 5; 41pp; English.
CC   Androgen receptor protein (AR) is used to produce mono- or poly-clonal
CC   antibodies. These are used for the detection and quantification of AR in
CC   the presence of endogenous androgen, as androgen will not interfere with
CC   binding. They may be used in assays to determine and quantify cellular
CC   distribution of AR in tumour tissue, and are esp. useful for evaluating
CC   prostate cancers to determine responsiveness to androgen withdrawal
CC   therapy.
SQ   Sequence   902 AA;

  Query Match            16.1%;  Score 570;  DB 1;  Length 902;
  Best Local Similarity  29.8%;  Pred. No. 1.48e-42;
  Matches  107;  Conservative  103;  Mismatches 127;  Indels  22;  Gaps  19;

Db     542 clicgdeasgchygaltcgsckvffkraaegkqkylcasrndctidkfrrkncpscrlrk 601
            | .|:| ||| |||   :| :|| ||||: :|    |:|:: |:||||| |||:| :||||
Qy      96 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 155

Db     602 cyeagmtl-garklk---klgnlklqeegenssagsptedpsqkmtvshi--egyecqpi 655
            ||| ||    |:|: :     :|          : :   || :     ::   |  :  ::   : :
Qy     156 CYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 215

Db     656 flnvleaiepgvvcaghdnnqpdsfaallsslnelgerqlvhvvkwakalpgfrnlhvdd 715
             | :|||  | |.  ...  .   | . |...  ||  |:::|||::  ||| :||| :| : |
Qy    216 VLTLLEAEPPHVL-ISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 273

Db    716 qmaviqyswmqlmvfamgwrsftnvnsrmlyfapdlvfn-eyrmhksrmysqcvrmrhls 774
          |: :::  || ::: :: |||. .   ...: ||||||:. .   .    .      .   :
Qy    274 QVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDMLLATT 331

Db    775 qetqwlqitpqeflcmkal-ll-fsiip-vdglknqkffdel-rm-nyike-ldriiack 828
            |  | :    |:||:||: ||   |. | |  .   .        | :. |  . . | :.||
Qy    332 SRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKS 391

Db    829 rknptscsrrfyqltklldsv-qpiarelhqftfdllikshmvsvdfpemmaeiisvqv 886
            ::   | |:  | :||  | :.   . .           :   | |         |  .  |
Qy    392 GISSQQQSMRLANLLMLLSHVRHASNKGMEHL-LNMKCK-NVVPV-Y-DLLLEMLNAHV 446

RESULT   11
ID   P91006 standard; protein; 902 BP.
AC   P91006;
DT   28-FEB-1990  (first entry)
DE   Rat androgen receptor DNA clone.
KW   Androgen receptor; TR2 polypeptide;
OS   Rat.
FH   Key              Location/Qualifiers
FT   Region           1..902
FT   /*tag-  a
FT   /product=98 kD polypeptide
FT   Region           170..902
FT   /*tag-  b
FT   /product=79 kD polypeptide
PN   WO8909223-A.
PD   05-OCT-1989.
PF   24-MAR-1989; U01238.
PR   30-MAR-1988; US-176107.
PA   (ARCH-) Arch Development Corp.
PI   Liao S, Chang C;
DR   WPI; 89-309501/42.
DR   N-PSDB; N91578.
PT   New DNA encoding new androgen receptor and TR2 polypeptide(s) - able
PT   to bind DNA, and derived antibodies, useful for receptor assay and
PT   purification.
PS   Claim 8; fig. 3; 60pp; English.
CC   The protein is used to raise antibodies for receptor assays and for
CC   affinity purification.
CC   The 98 kD product starts at the first Met codon, the 79 kD product
CC   starts from the second
SQ   Sequence   902 AA;

  Query Match            16.1%;  Score 570;  DB 1;  Length 902;
  Best Local Similarity  29.8%;  Pred. No. 1.48e-42;
  Matches  107;  Conservative  103;  Mismatches 127;  Indels  22;  Gaps  19;

Db    542 clicgdeasgchygaltcgsckvffkraaegkqkylcasrndctidkfrrknepscrlrk 601
          | :|:| ||| |||   :| :|| ||||. .|    |:|.  | ||||| ||| | .|||||
Qy     96 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 155

Db    602 cyeagmtl-garklk---klgniklqeegenssagsptedpsqkmtvshi--egyecqpi 655
          ||| ||    |:|: :    :|          . .     ||  .   ..   |  .  ..   . .
Qy    156 CYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 215

Db    656 flnvleaiepqvvcaqhdnnqpdsfaallsslnelgerqlvhvvkwakalpgfrnlhvdd 715
            | :|||  | |:  :. .  | . |:.  ||  |    :||:  ||   :||  : : |
Qy    216 VLTLLEAEPPHVL-ISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 273

Db    716 qmaviqyswmqlmvfamgwrsftnvnsrmlyfapdlvfn-eyrmhksrmysqcvrmrhls 774
          |: ::.  || ... .: |||. .   :... ||||||.. .   :     :      :   :
Qy    274 QVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDMLLATT 331

Db    775 qetqwlqitpqeflcmkal-ll-fsiip-vdglknqkffdel-rm-nyike-ldriiack 828
            |  |  :    |:||:||: ||   |. | |  .   .       | .. | : . | :.||
Qy    332 SRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKS 391

Db    829 rknptscsrrfyqltklldsv-qpiarelhqftfdllikshmvsvdfpemmaeiisvqv 886
            ::   | |:  | :||  | ::   : : :: :::    | ::|:| : ::: |::: :|

Qy    392 GISSQQQSMRLANLLMLLSHVRHASNKGMEHL-LNMKCK-NVVPV-Y-DLLLEMLNAHV 446

RESULT   12
ID   P12223 standard; Protein; 918 AA.
AC   R12223;
DT   20-AUG-1991  (first entry)
DE   Human androgen receptor.
KW   hAR; DNA-binding protein; steroid hormone.
OS   Homo sapiens.
FH   Key              Location/Qualifiers
FT   Domain           556..626
FT   /label= DNA binding domain
FT   /note= "cysteine-rich"
PN   WO9107423-A.
PD   30-MAY-1991.
PF   19-OCT-1990; U06015.
PR   17-NOV-1989; US-438775.
PA   (ARCH-) ARCH DEV CORP.
PI   Liao S, Chang C;
DR   WPI; 91-178048/24.
DR   N-PSDB; Q12001.
PT   Androgen receptor and TR2 DNA binding proteins - DNA sequences
PT   and antibodies for detection and quantification methods
PS   Claim 25; Fig 3; 79pp; English.
CC   This sequence was deduced from a cDNA clone isolated by screening
CC   commercially available human testis and prostate lambda gt11 cDNA
CC   libraries. The sequence is very similar to that of rat AR and in
CC   the DNA-binding domain it is identical to that of rAR DNA-binding
CC   domain. Homology comparisons with other known steroid receptors
CC   indicate that hAR is more closely related to glucocorticoid,
CC   mineralo-corticoid and progesterone receptors than to v-erb-A or to
CC   receptors for oestrogen, vitamin D and thyroid hormones.
SQ   Sequence   918 AA;

  Query Match            16.1%;  Score 570;  DB 3;  Length 918;
  Best Local Similarity  29.5%;  Pred. No. 1.48e-42;
  Matches  106;  Conservative  106;  Mismatches 125;  Indels  22;  Gaps  20;

Db    558 clicgdeasgchygaltcgsckvffkraaegkqkylcasrndctidkfrrknepscrlrk 617
          | .| | ||| |||    |  || ||||. .|    |:|:: |:||||| |||:| :|||||
Qy     96 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCIIDKNRRKSCQACRLRK 155

Db    618 cyeagmtl-garklk---klgn-iklqeeg-easstspteettqkltvshleqyecqpi 671
          ||| ||   | |      |      .|        . .       ..     :::    : :
Qy    156 CYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 215

Db    672 flnvleaiepgvvcaghdnnqpdsfaallsslnelgerqlvhvvkwakalpgfrnlhvdd 731
            : .|:|  | |:  .   .  | : |:.   :|  |     |:|    |||  .||| :| : |
Qy    216 VLTLLEAEPPHVL-ISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 273

Db    732 qmavicyswmglmvfamgwrsftnvnsrmlyfapdlvfn-eyrmhksrmysqcvrmrhls 790
          |  .  |  ||       .  |||. .   :  . ||||||            .       :   :
Qy    274 QVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVFGILEIFDMLLATT 331

Db    791 q-fgwlqitpqeflcmkal-ll-fsiip-vdglknqkffdel-rm-nyike-ldriiack 844
            |  |  .    |.||.||. ||   |: | |  :    :         | :: | : : | ::||
Qy    332 SRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKS 391

Db    845 rknptscsrrfyqltklldsv-qpiarelhqftfdllikshmvsvdfpemmaeiisvqv 902
                   | |   | :|: |    .  . :     ..   :    |.|  .  ...  |::: :|
Qy    392 GISSQQQSMRLANLLMLLSHVRHASNKGMEHL-LNMKCK-NVVPV-Y-DLLLEMLNAHV 446

RESULT   13
ID   P90996 standard; protein; 919 AA.
AC   P90996;
DT   28-FEB-1990  (first entry)
DE   Human androgen receptor DNA clone.
KW   Androgen receptor; TR2 polypeptide;
OS   Homo sapiens.
FH   Key              Location/Qualifiers
FT   Region           1..919
FT   /*tag=  a
FT   /product=98 kD polypeptide
FT   Region           185..919
FT   /*tag=  b
FT   /product=79 kD polypeptide
PN   WO8909223-A.
PD   05-OCT-1989.
PF   24-MAR-1989; U01238.
PR   30-MAR-1988; US-176107.
PA   (ARCH-) Arch Development Corp.
PI   Liao S, Chang C;
DR   WPI; 89-309501/42.
DR   N-PSDB; N91577.
PT   New DNA encoding new androgen receptor and TR2 polypeptide(s) - able
PT   to bind DNA, and derived antibodies, useful for receptor assay and
PT   purification.
PS   Claim 8; fig. 3; 60pp; English.
CC   The protein is used to raise antibodies for receptor assays and for
CC   affinity purification.
CC   The 98 kD product starts at the first Met codon; the 79 kD product
CC   starts from the second.
SQ   Sequence   919 AA.

  Query Match             16.1%;  Score 567;  DB 1;  Length 919;
  Best Local Similarity   29.2%;  Pred. No. 2.89e-42;
  Matches  105;  Conservative  107;  Mismatches  125;  Indels  22;  Gaps  20;

Db     559 clicgdeasgchygaltcgsckvffkraaegkqkylcasrndctidkfrrkncpscrlrk 618
           | .|.| ||| |||   .| .|| ||||. .|    |.|:: |.||||| |||:| .|||||
Qy      96 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 155

Db     619 cyeagmtl-garklk---klgn-lklqeeg-eassttspteettqkltvshiegyecqpi 672
           ||| ||   |:|: :     |       .|    ::      .       ..:    . :
Qy     156 CYEVCMVKCGSPPEPCGYPLVPRQPSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 215

Db     673 flnvleaiepgvvcaghdnnqpdsfaallsslnelgerqlvhvvkwakalpgfrnlhvdd 732
             | :|||  | |:  :: :  | . |     ||  |  .:|||:. ||| :||| :| : |
Qy     216 VLTLLEAEPPHVL-ISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 273

Db     733 qmaviqyswmglmvfamgwrsftnvnsrmlyfapdlvfn-eyrmhksrmysqcvrmrhls 791
           |: ::.  ||  .      |||.        |||||||:: :  .     :      :   :
Qy     274 QVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDMLLATT 331

Db     792 qefgwlqitpqeflcmkal-ll-fsiip-vdglknqkffdel-rm-nyike-ldriiack 845
              |   | :    |:||:||: ||   |: | | :    :         | :: | : : | ::||
Qy     332 SRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKS 391

Db     846 rknptscsrrfyqltklldsv-qpiarelhqftfdllikshmvsvdfpemmaeiisvqv 903
              ::    | |.   | .||   | :          : :      | :  | |  . :: :   .|
Qy     392 GISSQQQSMRLANLLMLLSHVPHASNKGMEHL-LNMKCK-NVVPV-Y-DLLLEMLNAHV 446

RESULT   14
ID   W14783 standard; Protein; 919 AA.
AC   W14783;
DT   22-JUN-1997  (first entry)
DE   Androgen receptor.
KW   Androgen receptor; acidic fibroblast growth factor; aFGF;
KW   antisense; benign prostatic hyperplasia, prostate cancer, therapy.
OS   Homo sapiens.
PN   WO9711170-A1.
PD   27-MAR-1997.
PF   20-SEP-1996; U15081.
PR   20-SEP-1995; US-004018.
PA   (WORC-) WORCESTER FOUND BIOMEDICAL RES
PI   Zamecnik PA,
DR   WPI; 97-201879/18.
DR   N-PSDB; T63407.
PT   Oligonucleotide(s) antisense to human androgen receptor and acidic
PT   FGF genes - used to inhibit gene expression, for the treatment of
PT   benign prostatic hyperplasia
PS   Disclosure; Page 22-28; 51pp; English.
CC   Human androgen receptor (W14783) binds testosterone and, acting
CC   at the transcriptional level, regulates the growth of normal
CC   prostatic cells.  Antisense oligonucleotides (see also T63200,
CC   T63404-05) based on an androgen receptor cDNA clone (see also
CC   T63407) can be used to prevent androgen receptor gene expression,
CC   thereby inhibiting the growth or survival of prostatic cells for
CC   the treatment of benign prostatic hyperplasia and prostate cancer.
SQ   Sequence   919 AA;

  Query Match             16.1%;  Score 567;  DB 21;  Length 919;
  Best Local Similarity   29.2%;  Pred. No. 2.89e-42;
  Matches  105;  Conservative  107,  Mismatches 125,  Indels  22;  Gaps  20;

Db     559 clicgdeasgchygaltcgsckvffkraaegkqkylcasrndctidkfrrkncpscrlrk 618
           | :|:| ||| |||  :| :|| ||||: :|    |:|:: |:||||| |||:| :|||||
Qy      96 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 155

Db     619 cyeagmtl-garklk---klgn-lklqeeg-eassttspteettqkltvshiegyecqpi 672
           ||| ||   |:|: .    .|     . :|    ..  .       .          .  : :
Qy     156 CYEVCMVKCGSPPEPCGYPLVPRQPSADEQLHCAGKAKRSGGHAPRVRFLLLDALSPEQL 215

Db     673 flnvleaiepgvvcaghdnnqpdsfaallsslnelgerqlvhvvkwakalpgfrnlhvdd 732
             | :|||  | |:  .. :  | . |.   ||  |: : |||.: ||| :||| :| : |
Qy     216 VLTLLEAEPPHVL-ISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 273

Db     733 qmaviqyswmglmvfamgwrsftnvnsrmlyfapdlvfn-eyrmhksrmysqcvrmrhls 791
           |  .    ||  .  .  .  |||       .. |||||| :  .   .    .     :   :
Qy     274 QVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDMLLATT 331

Db     792 qefgwlqitpqeflcmkal-ll-fsiip-vdglknqkffdel-rm-nyike-ldriiack 845
              |   | .    |:||:||: ||   |: | | :    :         | :: | : : | ::||
Qy     332 SRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKS 391

Db     846 rknptscsrrfyqltklldsv-qpiarelhqftfdllikshmvsvdfpemmaeiisvqv 903
              ::    | |:   | :||   | ..   . : :. :..    | ::|:| : ::: |::: :|
Qy     392 GISSQQQSMRLANLLMLLSHVPHASNKGMEHL-LNMKCK-NVVPV-Y-DLLLEMLNAHV 446

RESULT   15
ID   P93109 standard; protein; 919 AA.
AC   P93109;
DT   19-MAR-1990  (first entry)
DE   Human androgen receptor.
KW   Human androgen receptor; ploycional antibody; cancer.
OS   Homo sapiens.
PN   WO8909791-A.
PD   19-OCT-1989.
PF   13-APR-1989; U01548.
PP   14-APR-1988; US-182646
PA   (UYNC-) University of North Carolina
PI   French FS, Wilson EM, Joseph DR, Lubahn DB;
DR   WPI; 89-324206/44.
DR   N-PSDB; N91772.
PT   DNA encoding androgen receptor protein - useful for transforming
PT   eukaryotic hosts for protein expression and subsequent antibody prodn.
PS   Disclosure; Fig. 4; 41pp; English.
CC   Androgen receptor protein (AR) is used to produce mono- or poly-clonal
CC   antibodies.  These are used for the detection and quantification of AR in
CC   the presence of endogenous androgen, as androgen will not interfere with
CC   binding.  They may be used in assays to determine and quantify cellular
CC   distribution of AR in tumour tissue, and are esp. useful for evaluating
CC   prostate cancers to determine responsiveness to androgen withdrawal
CC   therapy
SQ   Sequence   919 AA;

  Query Match             16.0%;  Score 566;  DB 1;  Length 919;
  Best Local Similarity   29.2%;  Pred. No. 3.62e-42;
  Matches  105;  Conservative  107;  Mismatches 125,  Indels  22;  Gaps  20;

Db     559 clicgdeasgchygaltcgsckvffkraaegkqkylcasrndctidkfrrkncpscrlrk 618
           | :|:| ||| |||  :| :|| ||||: :|    |:|:: |:||||| |||:| :|||||
Qy      96 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 155

Db     619 cyeagmtl-garklk---klgn-lklqeeg-eassttspteettqkltvshiegyecqai 672
```

```
           ||| ||   |:|: :   :|    :  :|    :: :  :   : ::    :::   : :
Qy     156 CYEVGMVKCGSRRFRCGYRLVRRQRSADEQLHCAGKAKRSGCHAPRVRELLLDALSPEQL 215

Db     673 flnvleaiepqvvcaqhdnrqpdrfaallsslnelqerqlvhvvkwakalpqfrnlhvdd 732
            | :|||  | |:  :: :  | : |::: ||  |::::|||:: ||| :||| :| : |
Qy     216 VLTLLEAEPPHVL-ISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 273

Db     733 qmaviqyswmqlmvfamqwrsftnvnsrmlyfapdlvfn-eyrmhksrmysqcvrmrhls 791
           |: :::  || ::: :: |||: :  :::: ||||||:: :  :    :      :   :
Qy     274 QVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDMLLATT 331

Db     792 qefqwlqitpqeflcmkal-ll-fsiip-vdqlknqkffdel-rm-nyike-ldriiack 845
             |  | :   |:||:||: ||  |: | | :  :      | :: | : : | ::||
Qy     332 SRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKS 391

Db     846 rknptscsrrfyqltklldsv-qpiarelhqftfdllikshmvsvdfpcmmaelisvqv 903
             ::   | |:  | :||  | ::  : : :: :::  | ::|:| : ::: |::: :|
Qy     392 GISSQQQSMRLANLLMLLSHVRHASNKGMEHL-LNMKCK-NVVPV-Y-DLLLEMLNAHV 446

Search completed: Thu Dec 18 08:42:39 1997
Job time : 87 secs.
```

```
*************************************************************************

                         M P S R C H                               (TM)

*************************************************************************

          Release 2.1D John F. Collins, Biocomputing Research Unit.
          Copyright (c) 1993, 1994, 1995  University of Edinburgh, U K.
                    Distribution rights by IntelliGenetics, Inc.

MPsrch_pp   protein - protein database search, using Smith-Waterman algorithm

Run on.           Thu Dec 18 08:32:28 1997;  MasPar time 6.82 Seconds
                                             420.180 Million cell updates/sec

Tabular output not generated.

Title:            >US-08-826-361A-4
Description:      (1-233) from US08826361A.pep
Perfect Score:    1726
Sequence.         1 LVLTLLEAEPPHVLISRPSA  .......CKNVVPVYDLLLEMLNAHVL 233

Scoring table:    PAM 150
                  Gap 11

Searched:         101610 seqs, 12294212 residues

Post-processing:  Minimum Match 0%
                  Listing first 45 summaries

Database:         a-geneseq28
                    1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
                    8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
                    14:part14 15:part15 16:part16 17:part17 18:part18
                    19:part19 20:part20 21:part21

Statistics:       Mean 32.552,  Variance 149.049,  scale 0.218

       Pred. No. is the number of results predicted by chance to have a
       score greater than or equal to the score of the result being printed,
       and is derived by analysis of the total score distribution.

                                   SUMMARIES

                  %
Result          Query
  No.   Score   Match Length DB  ID             Description            Pred. No.
-------------------------------------------------------------------------------
   1     1726   100.0    485 21  W14724         Human oestrogen recep  1.25e-146
   2     1645    95.3    485 21  W14723         Rat oestrogen recepto  4.49e-139
   3     1625    94.1    485 21  W14725         Mouse oestrogen recep  3.29e-137
   4     1104    64.0    595  2  P70543         Human oestrogen recep  8.15e-89
   5      912    52.8    334  2  P80925         Sequence of the human  4.15e-71
   6      486    28.2    433  2  P80931         Sequence of human est  2.32e-32
   7      444    25.7    521  2  P80930         Sequence of human est  1.29e-28
   8      307    17.8    777  2  P80919         Sequence of the prima  1.38e-16
   9      299    17.3    510  8  R39469         hRXR-beta2             6.76e-16
  10      299    17.3    533  8  R39468         hRXR-beta1.            6.76e-16
  11      299    17.3    533 13  R72483         Human H-2RIIBP         6.76e-16
  12      296    17.1    446 13  R72484         H-2RIIBP               1.23e-15
  13      296    17.1    446  2  R11837         H2RIIBP nuclear hormo  1.23e-15
  14      292    16.9    448  8  R39467         mRXR-beta.             2.71e-15
  15      288    16.7    534 20  W03197         Glucocorticoid recept  5.97e-15
  16      286    16.6    467  3  R13799         Mouse RXR-alpha retin  8.86e-15
  17      286    16.6    467  8  R39470         mRXR-alpha.            8.86e-15
  18      287    16.6    984  2  P80927         Sequence of the human  7.27e-15
  19      283    16.4    463  8  R39471         mRXR-gamma.            1.60e-14
  20      282    16.3    462  3  R13798         Human RXR-alpha retin  1.95e-14
  21      282    16.3    462 20  W[illegible]   Human retinoid X recep 1.95e-14
  22      280    16.2    463  3  R13800         Mouse RXR-gamma retin  2.89e-14
  23      270    15.6    902  3  R12224         Rat androgen receptor  2.06e-13
  24      267    15.5    918  3  R12223         Human androgen recept  3.72e-13
  25      264    15.3    902  1  P93110         Rat androgen receptor  6.69e-13
  26      264    15.3    902  1  P91006         Rat androgen receptor  6.69e-13
  27      264    15.3    919  1  P93109         Human androgen recept  6.69e-13
  28      264    15.3    919  1  P90996         Human androgen recept  6.69e-13
  29      264    15.3    919 21  W14783         Androgen receptor.     6.69e-13
  30      248    14.4    630  3  R12230         TrpE/androgen recepto  1.51e-11
  31      211    12.2    558  7  R33746         XP5.                   1.86e-08
  32      202    11.7    458  2  R10549         Skin-specific murine   1.03e-07
  33      201    11.6    447  4  R20466         Mouse retinoic acid r  1.24e-07
  34      201    11.6    454 16  R84723         Human retinoic acid r  1.24e-07
  35      201    11.6    454  2  R10192         Recombinant human gam  1.24e-07
  36      201    11.6    454  2  R10550         Human Retinoic Acid R  1.24e-07
  37      201    11.6    458  4  R20465         Mouse retinoic acid r  1.24e-07
  38      201    11.6    458 16  R84726         Murine retinoic acid   1.24e-07
  39      196    11.4    598  9  R48631         Sequence of nuclear r  3.18e-07
  40      194    11.2    448 16  R84725         Murine retinoic acid   4.63e-07
  41      194    11.2    448  2  R10548         Murine Retinoic Acid   4.63e-07
  42      191    11.1     60 14  R78741         Glucocorticoid recept  8.12e-07
  43      190    11.0    448  1  P90341         hap (hepatoma) protei  9.79e-07
  44      189    11.0    448 10  R55128         Retinoic acid recepto  1.18e-06
  45      187    10.8    448 12  R66764         Human RAR-beta.        1.72e-06

                                   ALIGNMENTS

RESULT     1
ID    W14724 standard; Protein; 485 AA.
AC    W14724;
DT    08-JUN-1997  (first entry)
DE    Human oestrogen receptor beta (ER-beta).
KW    Orphan receptor; oestrogen receptor beta; ER-beta;
KW    nuclear receptor; prostate cancer; benign prostatic hyperplasia;
KW    ovary cancer; cardiovascular disease; osteoporosis;
KW    environment; pollutant.
OS    Homo sapiens.
FH    Key                Location/Qualifiers
FT    Domain             104..169
FT    /label= DNA-binding_domain
FT    Domain             260..457
FT    /label= Ligand-binding_domain
PN    WO9709348-A2.
PD    13-MAR-1997.
PF    09-SEP-1996; E03933.
PR    08-SEP-1995; GB-018272.
PR    15-MAR-1996; GB-005550.
PR    11-APR-1996; GB-007532.
PR    08-MAY-1996; GB-009576.
PA    (KARO-) KARO BIO AB.
PI    Enmark E,  Gustafsson J,  Kuiper GG;
DR    WPI; 97-192842/17.
DR    N-PSDB; T62843.
PT    New isolated oestrogen receptor beta - used to develop prods  for
PT    treating e.g. cancers, CNS diseases, osteoporosis or cardiovascular
PT    disease
PS    Claim 1; Fig 13A; 45pp, English.
CC    A novel human orphan receptor (W14724) is related to the known
CC    oestrogen receptor ER-alpha, and has been designated ER-beta.
CC    It is an oestrogen receptor-related nuclear receptor.  The ER-beta
CC    amino acid sequence was deduced from a cDNA clone (T62843) isolated
CC    from a human testis cDNA library.  Rat, human and mouse ER-beta
CC    (W14723-25) can be used to isolate molecules for use in the
CC    treatment of cardiovascular diseases, central nervous system
CC    diseases, osteoporosis, prostate or ovarian cancer or benign
CC    prostatic hyperplasia and to test environmental chemicals for
CC    oestrogenic activity.
SQ    Sequence   485 AA;

  Query Match            100.0%;  Score 1726;  DB 21;  Length 485;
  Best Local Similarity 100.0%;  Pred. No. 1.25e-146;
  Matches   233;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Db      223 lvltlleaepphvlisrpsapfteasmmmsltkladkelvhmiswakkipgfvelslfdq 282
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        1 LVLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFDQ 60

Db      283 vrllescwmevlmmglmwrsidhpgklifapdlvldrdegkcvegileifdmllattsrf 342
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       61 VRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSRF 120

Db      343 relklqhkeylcvkamillnssmyplvtatqdadssrklahllnavtdalvwviaksgis 402
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      121 RELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKSGIS 180

Db      403 sqqqsmrlanllmllshvrhasnkgmehllnmkcknvvpvydlllemlnahvl 455
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      181 SQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNAHVL 233


RESULT    2
ID   W14723 standard; Protein; 485 AA.
AC   W14723;
DT   08-JUN-1997  (first entry)
DE   Rat oestrogen receptor beta (ER-beta).
KW   Orphan receptor; oestrogen receptor beta; ER-beta;
KW   nuclear receptor; prostate cancer; benign prostatic hyperplasia;
KW   ovary cancer; cardiovascular disease; osteoporosis;
KW   environment; pollutant.
OS   Rattus sp.
FH   Key                 Location/Qualifiers
FT   Domain              104..169
FT   /label= DNA-binding_domain
FT   Domain              259..457
FT   /label= Ligand-binding_domain
PN   WO9709348-A2.
PD   13-MAR-1997.
PF   09-SEP-1996; F03933.
PR   08-SEP-1995; GB-018272.
PR   15-MAR-1996; GB-005550.
PR   11-APR-1996; GB-007532.
PR   08-MAY-1996; GB-009576.
PA   (KARO-) KARO BIO AB.
PI   Enmark E,  Gustafsson J,  Kuiper GG;
DR   WPI; 97-192842/17.
DR   N-PSDB; T62842.
PT   New isolated oestrogen receptor beta - used to develop prods. for
PT   treating e.g. cancers, CNS diseases, osteoporosis or cardiovascular
PT   disease
PS   Claim 1; Fig 1; 45pp; English.
CC   A novel rat orphan receptor (W14723) is related to the known
CC   oestrogen receptor ER-alpha, and has been designated ER-beta.
CC   It is an oestrogen receptor-related nuclear receptor.  The ER-beta
CC   amino acid sequence was deduced from a cDNA clone (T62842) isolated
CC   from a rat prostate cDNA library.  Rat, human and mouse ER-beta
CC   (W14723-25) can be used to isolate molecules for use in the
CC   treatment of cardiovascular diseases, central nervous system
CC   diseases, osteoporosis, prostate or ovarian cancer or benign
CC   prostatic hyperplasia and to test environmental chemicals for
CC   oestrogenic activity.
SQ   Sequence   485 AA;

  Query Match             95.3%;  Score 1645;  DB 21;  Length 485;
  Best Local Similarity   93.1%;  Pred. No. 4.49e-139;
  Matches   217;  Conservative   11;  Mismatches    5;  Indels    0;  Gaps    0;

Db      223 lvltlleaeppnvlvsrpsmpfteasmmmsltkladkelvhmigwakkipgfvelslldq 282
            |||||||||||:|:|||| |||||||||||||||||||||||||:|||||||||||||:||
Qy        1 LVLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFDQ 60

Db      283 vrllescwmevlmvglmwrsidhpgklifapdlvldrdegkcvegileifdmllattsrf 342
            |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Qy       61 VRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSRF 120

Db      343 relklqhkeylcvkamillnssmyplasanqeaessrklthllnavtdalvwviaksgis 402
            |||||||||||||||||||||||||| :| |:|:|||||:||||||||||||||||||||
Qy      121 RELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKSGIS 180

Db      403 sqqqsvrlanllmllshvrhisnkgmehllsmkcknvvpvydlllemlnahtl 455
            |||||:|||||||||||||| ||||||||:|||||||||||||||||||| |
Qy      181 SQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNAHVL 233


RESULT    3
ID   W14725 standard; Protein; 485 AA.
AC   W14725;
DT   08-JUN-1997  (first entry)
DE   Mouse oestrogen receptor beta (ER-beta).
KW   Orphan receptor; oestrogen receptor beta; ER-beta;
KW   nuclear receptor; prostate cancer; benign prostatic hyperplasia;
KW   ovary cancer; cardiovascular disease; osteoporosis;
KW   environment; pollutant.
OS   Mus sp.
FH   Key                 Location/Qualifiers
FT   Domain              104..169
FT   /label= DNA-binding_domain
FT   Domain              260..457
FT   /label= Ligand-binding_domain
PN   WO9709348-A2.
PD   13-MAR-1997.
PF   09-SEP-1996; F03933.
PR   08-SEP-1995; GB-018272.
PR   15-MAR-1996; GB-005550.
PR   11-APR-1996; GB-007532.
PR   08-MAY-1996; GB-009576.
PA   (KARO-) KARO BIO AB.
PI   Enmark E,  Gustafsson J,  Kuiper GG;
DR   WPI; 97-192842/17.
DR   N-PSDB; T62844.
PT   New isolated oestrogen receptor beta - used to develop prods. for
PT   treating e.g. cancers, CNS diseases, osteoporosis or cardiovascular
PT   disease
PS   Claim 1; Fig 14A; 45pp; English.
CC   A novel mouse orphan receptor (W14725) is related to the known
CC   oestrogen receptor ER-alpha, and has been designated ER beta.
CC   It is an oestrogen receptor-related nuclear receptor.  The ER-beta
CC   amino acid sequence was deduced from a cDNA clone (T62844).
CC   Rat, human and mouse ER beta (W14723-25) can be used to isolate
CC   molecules for use in the treatment of cardiovascular diseases,
CC   central nervous system diseases, osteoporosis, prostate or ovarian
CC   cancer or benign prostatic hyperplasia and to test environmental
CC   chemicals for oestrogenic activity.
SQ   Sequence   485 AA;

  Query Match             94.1%;  Score 1625;  DB 21;  Length 485;
  Best Local Similarity   91.8%;  Pred. No. 3.29e-137;
  Matches   214;  Conservative   13;  Mismatches    6;  Indels    0;  Gaps    0;

Db      223 lvltlleaeppnvlvsrpsmpfteasmmmsltkladkelvhmigwakkipgfvelslldq 282
            |||||||||||:|:|||| |||||||||||||||||||||||||:|||||||||||||:||
Qy        1 LVLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFDQ 60

Db      283 vrllescwmevlmvglmwrsidhpgklifapdlvldrdegkcvegileifdmllattarf 342
            |||||||||||||:|||||||||||||||||||||||||||||||||||||||||||:||
Qy       61 VRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSRF 120

Db      343 relklqhkeylcvkamillnssmyhlatasqeaessrklthllnavtdalvwvisksris 402
            |||||||||||||||||||||||| | ||:|:|:|||||:||||||||||||||:|| ||
Qy      121 RELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKSGIS 180

Db      403 sqqqsvrlanllmllshvrhisnkgmehllsmkcknvvpvydlllemlnahtl 455
            |||||:|||||||||||||| ||||||||:|||||||||||||||||||| |
Qy      181 SQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNAHVL 233


RESULT    4
ID   P70543 standard; protein; 595 AA.
```

```
AC   P70543;
DT   24-FEB-1991  (first entry)
DE   Human oestrogen receptor protein.
KW   Human oestrogen receptor protein; expression system; assay;
OS   Homo sapiens.
PN   WO8705049-A.
PD   27-AUG-1987.
PF   18-FEB-1987; U00341
PR   20-FEB-1986; US-833829
PA   (CALB-) California Biotechnology Inc.
PI   Shine J;
DR   WPI; 87-250213/35
DR   N-PSDB; N70880
PT   Expression system for vertebrate steroid receptor protein - comprising
PT   DNA sequence encoding the protein linked to control sequences in
PT   eucaryotic hosts
PS   Disclosure; Fig. 1-1 - 1-2; 26pp; English
CC   steroid receptor protein;
CC   The protein is expressed in an expression system from cDNA operably
CC   linked to control sequences compatible with eukaryotic host cells.
CC   This method allows expression under conditions which favour appropriate
CC   post-translational processing.  It also produces large amts. of purified
CC   protein useful in the design of agonist and antagonist cpds. for study
CC   of the mechanism of action of the steroid binding proteins in general,
CC   and for use in diagnostic assays for the proteins or antibodies to
CC   them.  These assays are important in, eg the diagnosis of tumour
CC   sensitivities to steroid metabolism  Suitable host cells are VERO,
CC   HeLa and CHO cells.
SQ   Sequence   595 AA;

  Query Match            64.0%;  Score 1104;  DB 2;  Length 595;
  Best Local Similarity  59.3%;  Pred  No  8.15e-89;
  Matches   140;  Conservative   61;  Mismatches  31;  Indels   4;  Gaps   3;

Db     315 mvsalldaepp-ilyseydptrpfseasmmglltnladrelvhminwakrvpgfvdltlh 373
            :| :||:|||| .| |     |: ||:||||| ||:|||:|||||:|||:·||||·|:|
Qy       1 LVLTLLEAEPPHVLISR--PSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLF 58

Db     374 dqvhllecawleilmiglvwrsmehpvkllfapnllldrnqgkcvegmveifdmllatss 433
           |||:|||  |:|:||:||·|||··|| ||.|||:|.||||:|||||||::|||||||||·|
Qy      59 DQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTS 118

Db     434 rfrmmnlqgeefvclksiillnsgvytflsstlksleekdhihrvldkitdtlihlmaka 493
           ||| ::||  |::|:|: ||||| :|  : :  '  .       .  .|· ·'|·|: ··'|:
Qy     119 RFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSRKLAHLLNAVTDALVWVIAKS 177

Db     494 rltlqqhqrlaqllllilshirhmsnkgmehlysmkcknvvplydlllemldahrl 549
             :: |||  ||| ||::|||:|| |||||||| :|||||||||·|||||||||:|| |
Qy     178 GISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNAHVL 233


RESULT    5
ID   P80925 standard; protein; 334 AA.
AC   P80925;
DT   29-DEC-1990  (first entry)
DE   Sequence of the human oestrogen receptor
KW   Hormone receptor, hormone-binding, transcription activation;.
OS   Homo sapiens.
PN   WO8803168-A.
PD   05-MAY-1988.
PF   23-OCT-1987; U02782.
PR   20-OCT-1987; US-108471
PA   (SALK) Salk Inst for Biol Stud.
PI   Evans RM, Weinberger CA, Hollenberg SM, Giguere V:
DR   WPI; 88-133242/19.
PT   Recombinant DNA encoding hormone recpetors -
PT   comprising glucocorticoid, mineralocorticoid, thyroid hormone
PT   and novel hormone receptors
PS   Example; Fig III-2; 243pp; English.
CC   DNAs encoding hormone receptors and the hormone receptors themselves are
CC   claimed  The DNA can be used to make the hormone receptor proteins and
CC   functional modified forms in quantities not previously possible. The
CC   receptor proteins can be used to screen cpds. for receptor-agonist or
CC   receptor-antagonist activity. They can also be used in diagnostic assays.
SQ   Sequence   334 AA;

  Query Match            52.8%;  Score 912;  DB 2;  Length 334;
  Best Local Similarity  56.1%;  Pred  No  4.15e-71;
  Matches   115;  Conservative   58;  Mismatches  28;  Indels   4;  Gaps   3;

Db     131 mvsalldaepp-ilyseydptrpfseasmmglltnladrelvhminwakrvpgfvdltlh 189
            .| .||.|||| .| |     |. ||.||||| ||·|||:|||||||·|||··||||·|·|
Qy       1 LVLILLEAEPPHVLISR--PSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLF 58

Db     190 dqvhllecawleilmiglvwrsmehpvkllfapnllldrnqgkcvegmveifdmllatss 249
           |||.||:|  |:|:||·|||:||::.|| ||.|||:|·:|||·:||||||:··||||||||·|
Qy      59 DQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTS 118

Db     250 rfrmmniqgeefvclksiillnsgvytflsstlksleekdhihrvldkitdtlihlmaka 309
           ||| :.|| |.:|:|..||||||:·| ::··|   · ·     · ·:|: :||:|: ::||:
Qy     119 RFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSRKLAHLLNAVTDALVWVIAKS 177

Db     310 gltlqqhqrlaqllllilshirhms 334
           |:: |||  ||| ||::|||:|| |
Qy     178 GISSQQQSMRLANLLMLLSHVRHAS 202


RESULT    6
ID   P80931 standard, protein; 433 AA.
AC   P80931;
DT   29-DEC-1990  (first entry)
DE   Sequence of human estrogen-related receptor protein (hERR2)
KW   Hormone receptor; hormone-binding; transcription activation.
OS   Homo sapiens.
PN   WO8803168-A.
PD   05-MAY-1988.
PF   23-OCT-1987; U02782.
PR   20-OCT-1987; US-108471.
PA   (SALK) Salk Inst for Biol Stud.
PI   Evans RM, Weinberger CA, Hollenberg SM, Giguere V,
DR   WPI; 88-133242/19.
DR   N-PSDB; N80922.
PT   Recombinant DNA encoding hormone recpetors -
PT   comprising glucocorticoid, mineralocorticoid, thyroid hormone
PT   and novel hormone receptors
PS   Claim 17, Fig V-2(B)-1 and -2; 243pp; English
CC   DNAs encoding hormone receptors and the hormone receptors themselves are
CC   claimed. The DNA can be used to make the hormone receptor proteins and
CC   functional modified forms in quantities not previously possible. The
CC   receptor proteins can be used to screen cpds. for receptor-agonist or
CC   receptor-antagonist activity. They can also be used in diagnostic assays.
CC   Also claimed is pure DNA encoding oestrogen-related receptors hERR1 and
CC   hERR2 and hERR1 and hERR2. The new hERR1 and hERR2 receptors will provide
CC   the basis for development of an assay system that will lead to the
CC   identification of novel hormones.
SQ   Sequence   433 AA;

  Query Match            28.2%;  Score 486;  DB 2;  Length 433;
  Best Local Similarity  36.1%;  Pred. No  2.32e-32;
  Matches    83;  Conservative   63;  Mismatches  70;  Indels  14;  Gaps   8;

Db     216 llvacpdklyamppddvpegdikalttlcdladrelvfliswakhipgfsnltlgdqmsl 275
           || |||  :     |    |   :     : :|  |||:|||| :||||| |||| .|:| |||: |
Qy       5 LLEAEPPHVLISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFDQVRL 63

Db     276 lqsawmeililgivyrslpyddklayaedyimdeehsrlv-gllelyrailqlvrrykkl 334
           |.| |||.|::|:: ||:      || :| | :·| : :: | |:||:: :|    |:: |
Qy      64 LESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSRFREL 123

Db     335 kvekeefvmlkalalansdsmy--i---enleavqklqdllh--ealq-dycls--qrh 383
           |.. |:: .||. | | | ||          ||  ||·   ·||         |   :
Qy     124 KLQHKEYLCVKAMILLNS-SMYPLVTATQDADSSPKLAHLLNAVTDALVWVIAKSGISSQ 182

Db     384 eeprragklltlpllrqtaakavqhfysvklqgkvpmhklflemleakv 433
           ... | .:|.. |.  ·  .    . .  .|     |·.  '·||||:| |
Qy     183 QQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNAHV 232
```

```
RESULT     7
ID   P80930 standard; protein; 521 AA
AC   P80930;
DT   10-MAR-1993  (revised)
DT   29-DEC-1990  (first entry)
DE   Sequence of human estrogen-related receptor protein (hERR1)
KW   Hormone receptor; hormone binding; transcription activation
OS   Homo sapiens.
PN   WO8803168-A.
PD   05-MAY-1988.
PF   23-OCT-1987; U02782.
PR   20-OCT-1987; US-108471.
PA   (SALK) Salk Inst for Biol Stud
PI   Evans RM, Weinberger CA, Hollenberg SM, Giguere V;
DR   WPI; 88-133242/19
DR   N-PSDB; N80921.
PT   Recombinant DNA encoding hormone recepetors -
PT   comprising glucocorticoid, mineralocorticoid, thyroid hormone
PT   and novel hormone receptors
PS   Claim 17; Fig V-1(B)-1 and -2; 243pp; English.
CC   DNAs encoding hormone receptors and the hormone receptors themselves are
CC   claimed. The DNA can be used to make the hormone receptor proteins and
CC   functional modified forms in quantities not previously possible. The
CC   receptor proteins can be used to screen cpds  for receptor-agonist or
CC   receptor antagonist activity. They can also be used in diagnostic assays
CC   Also claimed is pure DNA encoding oestrogen-related receptors hERR1 and
CC   hERR2 and hERR1 and hERR2. The new hERR1 and hERR2 receptors will provide
CC   the basis for development of an assay system that will lead to the
CC   identification of novel hormones
SQ   Sequence   521 AA;

  Query Match             25.7%;  Score 444;  DB 2;  Length 521;
  Best Local Similarity   36.4%;  Pred. No. 1.29e-28;
  Matches    84;  Conservative   51;  Mismatches  89;  Indels   7;  Gaps   4;

Db      294 lvshllvvepeklyampdpagpdghlpavatlcdlfdreivvtiswaksipgfsslslsd 353
            ||  ||  ||  :  :: |::|      : : :|  | |:|:|  ||||| ||||  ||| |
Qy        1 LVLTLLEAEPPHVL-ISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 59

Db      354 qmsvlqsvwmevlvlgvaqrslplqdelafaedlvld-eegaraaglgelgaallqlvrr 412
            |: :|:| ||||::|:  ||:        | || ||||| :||      |: |:   ||     |
Qy       60 QVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSR 119

Db      413 lqalrlereeyvllkalalanedsvhiedeprlwsseekllh--eal---leyeagragp 467
            :: |:|:: ||: :||: | ||       :      :    |  || |  :|:    |     :::|
Qy      120 FRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKSGI 179

Db      468 qqqaerrradrllltlpllrqtagkvlahfygvklegkvpmhklflemlea 518
            ::   :   |  : ||: |: :|::: |  : |:   :|      ||:  |:||||:|
Qy      180 SSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNA 230


RESULT     8
ID   P80919 standard; protein; 777 AA
AC   P80919;
DT   29-DEC-1990  (first entry)
DE   Sequence of the primary protein sequence of human glucocorticoid
DE   receptor (hGR)
KW   Hormone receptor; hormone binding; transcription activation
OS   Homo sapiens.
PN   WO8803168-A.
PD   05-MAY-1988.
PF   23-OCT-1987; U02782.
PR   20-OCT-1987; US-108471.
PA   (SALK) Salk Inst for Biol Stud
PI   Evans RM, Weinberger CA, Hollenberg SM, Giguere V;
DR   WPI; 88-133242/19.
DR   N-PSDB; N80916.
PT   Recombinant DNA encoding hormone recepetors -
PT   comprising glucocorticoid, mineralocorticoid, thyroid hormone
PT   and novel hormone receptors
PS   Claim 24; Fig I-2(1) to (2); 243pp; English.
CC   DNAs encoding hormone receptors and the hormone receptors themselves are
CC   claimed. The DNA can be used to make the hormone receptor proteins and
CC   functional modified forms in quantities not previously possible. The
CC   receptor proteins can be used to screen cpds  for receptor-agonist or
CC   receptor-antagonist activity. They can also be used in diagnostic assays.
SQ   Sequence   777 AA;

  Query Match             17.8%;  Score 307;  DB 2;  Length 777;
  Best Local Similarity   32.4%;  Pred. No. 1.38e-16;
  Matches    47;  Conservative   46;  Mismatches  46;  Indels   6;  Gaps   5;

Db      532 lv-sllevlepevlyagydssvpdstwrimttlnmlggrqviaavkwakaipgfrnlhld 590
            || :|||     | ||  :    ::  :::   :| :| :|: ::::   : ||| |||| :| |
Qy        1 LVLTLLEAEPPHVLISRPSAPFTEAS--MMMSLTKLADKELVHMISWAKKIPGFVELSLF 58

Db      591 dqmtllqyswmflmafalgwrsyrqssanllcfapdliin-eqrmtlpcmydqckhmlyv 649
            ||: ||:  ||  ::   :| |||   ::: :|: |||||::: :: :  :   :  :     :|
Qy       59 DQVRLLESCWMEVLMMGLMWRSIDHPG-KLI-FAPDLVLDRDEGKCVEGILEIFDMLLAT 116

Db      650 sselhrlqvsyeeylcmktllllss 674
            :|  :: |  :    ||||:|:::||:|
Qy      117 TSRFRELKLQHKEYLCVKAMILLNS 141


RESULT     9
ID   R39469 standard; Protein; 510 AA.
AC   R39469;
DT   01-FEB-1994  (first entry)
DE   hRXR-beta2.
KW   Retinoid X receptor; RXR; retinoic acid receptor; RAR;
KW   thyroid receptor; TR; retinoic acid response elements; RARE;
KW   thyroid receptor response element; TRE;
KW   retinoid X receptor response element; RXRE;
KW   subunit; heterodimer.
OS   Homo sapiens.
PN   WO9315216-A.
PD   05-AUG-1993.
PF   25-JAN-1993; U00639.
PR   24-JAN-1992; US-825667.
PA   (CNRS ) CENT NAT RECH SCI.
PA   (INRM ) INST NAT SANTE & RECH MEDICALE.
PA   (SQUI ) SQUIBB & SONS INC E R.
PA   (UYST-) UNIV PASTEUR STRASBOURG LOUIS.
PI   Chambon P,  Kastner P,  Leid M;
DR   WPI; 93-259691/32.
DR   N-PSDB; Q45957.
PT   Heterodimer comprising retinoic acid or thyroid receptor and
PT   retinoid X receptor - useful in diagnosis and for generation of
PT   antibodies
PS   Claim 4; Page 92-94; 128pp; English.
CC   Two different types of nuclear receptors, RAR and TR dimerise with
CC   RXR to form a heterodimer which is capable of binding to RARE, TRE
CC   or RXRE at physiological conditions.  Sequences encoding 1 isoform
CC   of mouse RXR-beta, 3 isoforms of human RXR beta, 1 isoform of mouse
CC   RXR-alpha and 1 isoform of mouse RXR-gamma are provided ([illegible]).
SQ   Sequence   510 AA;

  Query Match             17.3%;  Score 299;  DB 8;  Length 510;
  Best Local Similarity   32.8%;  Pred. No. 6.76e-16;
  Matches    66;  Conservative   52;  Mismatches  72;  Indels  [illegible];  Gaps  10;

Db      320 adkqlftlvewakriphfsslplddqvillragwnelliasfshrsidvrdgillatglh 379
            |||:|   :: |||:||  |   |:| ||| ||  :  | |:|: ::  ||||      :::|  |
Qy       35 ADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLV 94

Db      380 vhrnsahsa-gvgaifdrvltelvskmrdmrmdktelgclraiilfnpdakglsnps-ev 437
            . |. .      |.  ||| .|.    |. |.::::  |   |::|:||:|:      |   :: :
Qy       95 LDRDEGKCVEGILEIFDMLLAT-TSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA 153

Db      438 ev---l-r--ekvyasle-tyckqkyp-eqqg-rfaklllrlpalrsiglkclehlfffk 488
            :      |  :   :  |   :|       |     :  :||: |:|:||:  |:  :|    : | :|||:    |
Qy      154 DSSRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMK 213
```

```
Db     489 ligdtpidtflmemleaphql 509
                   |:  :|:||||:| | |
Qy     214 CKNVVPVYDLLLEMLNA-HVL 233


RESULT   10
ID   R39468 standard; Protein; 533 AA.
AC   R39468;
DT   01-FEB-1994  (first entry)
DE   hRXR-beta1.
KW   Retinoid X receptor; RXR; retinoic acid receptor; RAR;
KW   thyroid receptor; TR; retinoic acid response elements; RARE;
KW   thyroid receptor response element; TRE;
KW   retinoid X receptor response element; RXRE;
KW   subunit; heterodimer.
OS   Homo sapiens.
PN   WO9315216-A.
PD   05-AUG-1993
PF   25-JAN-1993; U00639.
PR   24-JAN-1992; US-825667.
PA   (CNRS ) CENT NAT RECH SCI.
PA   (INRM ) INST NAT SANTE & RECH MEDICALE.
PA   (SQUI ) SQUIBB & SONS INC E R.
PA   (UYST-) UNIV PASTEUR STRASBOURG LOUIS
PI   Chambon P,  Kastner P,  Leid M;
DR   WPI; 93-258691/32
DR   N-PSDB; Q45956.
PT   Heterodimer comprising retinoic acid or thyroid receptor and
PT   retinoid X receptor - useful in diagnosis and for generation of
PT   antibodies
PS   Claim 4; Page 87-90; 128pp; English.
CC   Two different types of nuclear receptors, RAR and TR dimerise with
CC   RXR to form a heterodimer which is capable of binding to RARE, TRE
CC   or RXRE at physiological conditions.  Sequences encoding 1 isoform
CC   of mouse RXR-beta, 3 isoforms of human RXR-beta, 1 isoform of mouse
CC   RXR-alpha and 1 isoform of mouse RXR-gamma are provided (Q45955-60).
SQ   Sequence   533 AA;

  Query Match             17.3%;  Score 299;  DB 8;  Length 533;
  Best Local Similarity   32.8%;  Pred. No. 6.76e-16;
  Matches   66;  Conservative   52;  Mismatches  70;  Indels  13;  Gaps  10;

Db     343 adkqlftlvewakriphfsslplddqvillragwnelliasfshrsidvrdgillatglh 402
           |||:|  :: |||:|| |   |:| ||| || : | |:|: ::  ||||     :::|  :
Qy      35 ADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLV 94

Db     403 vhrnsahsa-gvgaifdrvltelvskmrdmrmdktelgclrailifnpdakglsnps-ev 460
           : |: :     |:  ||| :|.     |: |.....  |  |..|.||.|.     |  :: :
Qy      95 LDRDEGKCVEGILEIFDMLLAT-TSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA 153

Db     461 ev---l-r--ekvyasle-tyckqkyp-eqqg-rfakllrlpalrsiglkclehlfffk 511
           :    | :  : |  :|      |    : :||: |:|:||: |: :|    : | :|||:  |
Qy     154 DSSRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMK 213

Db     512 ligdtpidtflmemleaphql 532
                   |:  :|:||||:| | |
Qy     214 CKNVVPVYDLLLEMLNA-HVL 233


RESULT   11
ID   R72483 standard; Protein; 533 AA.
AC   R72483;
DT   24-OCT-1995  (first entry)
DE   Human H-2RIIBP.
KW   H-2 region II binding protein; H-2RIIBP, retinoid receptor.
OS   Homo sapiens.
PN   US5403925-A.
PD   04-APR-1995.
PF   13-DEC-1989; 450162.
PR   13-DEC-1989; US-450162.
PR   09-APR-1992; US-866950.
PR   28-SEP-1992; US-952800.
```

```
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PI   Ozato K;
DR   WPI; 95-146856/19.
DR   N-PSDB; Q89531.
PT   New DNA encoding human H-2 region II binding protein - an
PT   intracellular retinoid receptor mediating effects of steroid(s)
PT   etc., also diagnostic kits contg. probes derived from the DNA
PS   Claim 1; Column 15-20; 14pp; English.
CC   A cDNA library of the Burkitt lymphoma cell line Daudi was screened
CC   with a fragment of the murine retinoid X receptor-beta (H-2RIIBP)
CC   gene.  A positive clone contained an insert (sequence given in
CC   Q89531) that encoded human H-2RIIBP (R72483).  Recombinant human H-
CC   2RIIBP was expressed in Sf9 cells and used to raise antibodies in
CC   mice.
SQ   Sequence   533 AA;

  Query Match             17.3%;  Score 299;  DB 13;  Length 533;
  Best Local Similarity   32.8%;  Pred. No. 6.76e-16;
  Matches   66;  Conservative   52;  Mismatches  70;  Indels  13;  Gaps  10;

Db     343 adkqlftlvewakriphfsslplddqvillragwnelliasfshrsidvrdgillatglh 402
           |||:|  :: |||:|| |   |:| ||| || : | |:|: ::  ||||     :::|  :
Qy      35 ADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLV 94

Db     403 vhrnsahsa-gvgaifdrvltelvskmrdmrmdktelgclrailifnpdakglsnps-ev 460
           : |: :     |:  ||| :|:     |: |:::::  |  |::|:||:|:     |  :: :
Qy      95 LDRDEGKCVEGILEIFDMLLAT-TSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA 153

Db     461 ev---l-r--ekvyasle-tyckqkyp-eqqg-rfakllrlpalrsiglkclehlfffk 511
           :    |      |  |     |      ||: |:|:||: |: :|   :    :|||:  |
Qy     154 DSSRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMK 213

Db     512 ligdtpidtflmemleaphql 532
                   |:  :|:||||:| | |
Qy     214 CKNVVPVYDLLLEMLNA-HVL 233


RESULT   12
ID   R72484 standard; Protein; 446 AA.
AC   R72484;
DT   24-OCT-1995  (first entry)
DE   H-2RIIBP.
KW   H-2 region II binding protein, H-2RIIBP, retinoid receptor.
OS   Not specified.
PN   US5403925-A.
PD   04-APR-1995.
PF   13-DEC-1989; 450162.
PR   13-DEC-1989; US-450162.
PR   09-APR-1992; US-866950.
PR   28-SEP-1992; US-952800.
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PI   Ozato K;
DR   WPI; 95-146856/19.
DR   N-PSDB; Q89532.
PT   New DNA encoding human H-2 region II binding protein - an
PT   intracellular retinoid receptor mediating effects of steroid(s)
PT   etc., also diagnostic kits contg. probes derived from the DNA
PS   Disclosure; Column 23-26; 14pp; English.
CC   A cDNA library of the Burkitt lymphoma cell line Daudi was screened
CC   with a fragment of the murine retinoid X receptor-beta (H-2RIIBP)
CC   gene.  A positive clone contained an insert (sequence given in
CC   Q89531) that encoded human H-2RIIBP (R72483).  Further H-2RIIBP
CC   sequences of unspecified origin (possibly murine) are given in
CC   Q89432 and R72484.
SQ   Sequence   446 AA;

  Query Match             17.1%;  Score 296;  DB 13;  Length 446;
  Best Local Similarity   32.8%;  Pred. No. 1.23e-15;
  Matches   66;  Conservative   51;  Mismatches  71;  Indels  13;  Gaps  10;

Db     256 adkqlftlvewakriphfsslplddqvillragwnelliasfshrsidvrdgillatglh 315
           |||.|  :: |||:|| |   |:| ||| || : | |:|: ::  ||||     :::|  |
Qy      35 ADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLV 94
```

```
Db    316 vhrnsahsa-qvqaifdrvltelvskmrdmrmdktelqclraiilfnpdakql-snpgev 373
          : |: :      |:  ||| :|:    |: |:::::  |   |::|:||:|:      |  :     :
Qy     95 LDRDEGKCVEGILEIFDMLLAT-ISRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA 153

Db    374 ei---l-r--ekvyaslo-tyckqkyp-eqqg-rfaklllrlpalrsiglkclehlfffk 424
          :     |  :   :  |   :|        |     :  :||: |:|:||: |: :|    :  |  :|||:   |
Qy    154 DSSRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMK 213

Db    425 liqdtpidttlmemleaphql 445
                |:   :|:|||:|  |  |
Qy    214 CKNVVPVYDLLLEMLNA-HVL 233


RESULT   13
ID   R11837 standard; Protein; 446 AA.
AC   R11837;
DT   16-JUL-1991  (first entry)
DE   H2RIIBP nuclear hormone receptor protein
KW   Major histocompatibility complex, MHC class I regulatory element,
KW   nuclear hormone receptors; tumours; cancer; H2RIIBP.
OS   synthetic.
PN   US7450162-A.
PD   08-MAY-1990.
PF   13-DEC-1989; 163312.
PR   13-DEC-1989; US-450162.
PA   (USSH ) NAT INST OF HEALTH.
PI   Ozato K;
DR   WPI; 91-140556/19.
DR   N-PSDB; Q11675.
PT   New protein of nuclear hormone receptor super-family - used in
PT   diagnosis and prognosis of treatment of tumours and other
PT   conditions
PS   Disclosure; fig 3; 25pp; English.
CC   This H2RIIBP major histocompatibility complex (MHC) class I regul-
CC   atory element (CRE) binding factor binds to the wild-type MHC CRE
CC   region II sequence specifically. It does not bind any other
CC   sequences except for the Vitellogenin A2 ERE sequence. It can be
CC   used for the diagnosis and prognosis of tumours or conditions
CC   resulting from the imbalance in the normal level of H2RIIBP in
CC   cells or tissues.
CC   See also Q11664-74.
SQ   Sequence   446 AA;

  Query Match              17.1%;  Score 296;  DB 2;  Length 446;
  Best Local Similarity  32.8%;  Pred. No. 1.23e-15;
  Matches    66;  Conservative   51;  Mismatches  71;  Indels  13;  Gaps  10;

Db    256 adkqlftlvewakriphfsslplddqvillraqwnelliasfshrsidvrdqillatqlh 315
          |||:|   :: |||:|| |    |:| ||| ||  : | |:|: ::   ||||      :::|  |
Qy     35 ADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLV 94

Db    316 vhrnsahsa-qvqaifdrvltelvskmrdmrmdktelqclraiilfnpdakql-snpgev 373
          : |: :      |:  ||| :|:    |: |:::::  |   |::|:||:|:      |  :     :
Qy     95 LDRDEGKCVEGILEIFDMLLAT-ISRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA 153

Db    374 ei---l-r--ekvyaslo-tyckqkyp-eqqg-rfaklllrlpalrsiglkclehlfffk 424
          :     |  :   :  |   :|        |     :  :||: |:|:||: |: :|    :  |  :|||:   |
Qy    154 DSSRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMK 213

Db    425 liqdtpidttlmemleaphql 445
                |:   :|:|||:|  |  |
Qy    214 CKNVVPVYDLLLEMLNA-HVL 233


RESULT   14
ID   R39467 standard; Protein; 448 AA.
AC   R39467;
DT   01-FEB-1994  (first entry)
DE   mRXR-beta.
KW   Retinoid X receptor; RXR; retinoic acid receptor; RAR;
KW   thyroid receptor, TR, retinoic acid response elements, RARE,
KW   thyroid receptor response element, TRE,
```

```
KW   retinoid X receptor response element, RXRE, subunit, heterodimer.
OS   Mus musculus.
PN   WO9315216-A.
PD   05-AUG-1993.
PF   25-JAN-1993; U00639.
PR   24-JAN-1992; US-825667.
PA   (CNRS ) CENT NAT RECH SCI.
PA   (INRM ) INST NAT SANTE & RECH MEDICALE.
PA   (SQUI ) SQUIBB & SONS INC E R.
PA   (UYST-) UNIV PASTEUR STRASBOURG LOUIS.
PI   Chambon P,  Kastner P,  Leid M;
DR   WPI, 93-258691/32.
DR   N-PSDB; Q45955.
PT   Heterodimer comprising retinoic acid or thyroid receptor and
PT   retinoid X receptor - useful in diagnosis and for generation of
PT   antibodies
PS   Claim 4; Page 84-86; 128pp; English.
CC   Two different types of nuclear receptors, RAR and TR dimerise with
CC   RXR to form a heterodimer which is capable of binding to RARE, TRE
CC   or RXRE at physiological conditions.  Sequences encoding 1 isoform
CC   of mouse RXR-beta, 3 isoforms of human RXR-beta, 1 isoform of mouse
CC   RXR-alpha and 1 isoform of mouse RXR-gamma are provided (Q45955-60).
SQ   Sequence   448 AA;

  Query Match              16.9%;  Score 292;  DB 8;  Length 448;
  Best Local Similarity  32.3%;  Pred. No. 2.71e-15;
  Matches    65;  Conservative   52;  Mismatches  71;  Indels  13;  Gaps  10;

Db    258 adkqlftlvewakriphfsslplddqvillraqwnelliasfshrsidvrdqillatqlh 317
          |||:|   :: |||:|| |    |:| ||| ||  : | |:|: ::   ||||      :::|  |
Qy     35 ADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLV 94

Db    318 vhrnsahsa-qvqaifdrvltelvskmrdmrmdktelqclraiimfnpdakql-snpqev 375
          : |: :      |:  ||| :|:    |: |:::::  |   |::|:|::|:      |  :     :
Qy     95 LDRDEGKCVEGILEIFDMLLAT-ISRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA 153

Db    376 ei---l-r--ekvyaslo-tyckqkyp-eqqg-rfaklllrlpalrsiglkclehlfffk 426
          :     |  :   :  |   :|        |     :  :||: |:|:||: |: :|    :  |  :|||:   |
Qy    154 DSSRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMK 213

Db    427 ligdtpidtflmemleaphql 447
                |:   :|:|||:|  |  |
Qy    214 CKNVVPVYDLLLEMLNA-HVL 233


RESULT   15
ID   W03197 standard; Protein; 534 AA.
AC   W03197;
DT   16-APR-1997  (first entry)
DE   Glucocorticoid receptor/adenoviral ORF6 chimaeric fusion protein.
KW   Human adenovirus, inverted tandem repeat, E4, E1; open reading frame;
KW   expression, late gene, cell line, defective recombinant adenovirus;
KW   complementation, gene therapy, cystic fibrosis, dystrophy; PCR;
KW   neurodegenerative disease; blood coagulation disorder; viral infection;
KW   primer; amplification; polymerase chain reaction; fusion protein.
OS   Homo sapiens/Adenovirus Ad5.
PN   WO9622378-A1.
PD   25-JUL-1996.
PF   19-JAN-1996; F00088.
PR   20-JAN-1995; FR-000747.
PR   01-JUN-1995; FR-006532.
PR   08-SEP-1995; FR-010541.
PA   (RHON ) RHONE POULENC RORER SA.
PI   Dedieu J,  Latta M,  Orsini C,  Perricaudet M,  Vigne E;
PI   Yeh P,  Dedieu JF;
DR   WPI; 96-354535/35.
DR   N-PSDB; T31320.
PT   Cells for prodn. of recombinant adeno and adeno-associated virus -
PT   contain part of the E4 region of adeno virus, used to produce new
PT   adeno virus defective in E1 and E4 regions for use as gene therapy
PT   vectors
PS   Example 1, Page 41-43, 69pp; French.
CC   The invention relates to novel cell lines which can support the
CC   replication and production of defective recombinant adenoviruses which
CC   lack all or part of the E1 and E4 regions.  The novel cell lines contain
CC   the adenoviral E4 region, esp. the open reading frame 6 and/or 7,
CC   inserted into its gene under control of a promoter, pref. from MMTV which
CC   is regulated by dexamethasone.  The E4 ORF6 may also be inserted into the
CC   genome as a chimaeric fusion gene comprising a glucocorticoid receptor
CC   hormone binding domain fused N-terminally to the adenoviral ORF6 gene.
CC   The fusion protein generated (this sequence) is retained in the cytoplasm
CC   of the cell in the absence of hormone by interaction of the hormone
CC   binding domain with the hsp90 protein or other factors.  Contacting the
CC   fusion protein with the hormone causes translocation of the protein to
CC   the cell's nucleus were the ORF6 product functions.  The gene was
CC   generated by amplifying the sequence encoding the hormone binding domain
CC   with primers T31316-7 and recombining the fragment with the adenoviral
CC   ORF6 sequence.  The defective viruses generated in the new cells can be
CC   used for gene therapy e.g. for cystic fibrosis, dystrophy,
CC   neurodegenerative diseases, blood coagulation disorders, viral
CC   infections, etc..
SQ   Sequence   534 AA;

  Query Match            16.7%;  Score 288;  DB 20;  Length 534;
  Best Local Similarity  33.3%;  Pred. No. 5.97e-15;
  Matches   39;  Conservative   39;  Mismatches  36;  Indels   3;  Gaps   3;

Db      23 imttlnmlggrqviaavkwakaipgfrnlhlddqmtllqyswmflmafalgwrsyrqssa 82
              :| :| :|: ::::  : |||  |||| :| | ||: ||:  || ··  ·| |||  :::
Qy      27 MMMSLTKLADKFLVHMISWAKKIPGFVEISLFDQVPLIFSCWMEVLMMGLMWRSIDHPG- 85

Db      83 nllcfapdliin-eqrmtlpcmydqckhmlyvsselhrlqvsyeeylcmktllllss 138
              :|: |:|||::: :: : :  : :    :|  :| :: | :   ||||:|:::||:|
Qy      86 KLI-FAPDIVIPPDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNS 141

Search completed. Thu Dec 18 08:33:45 1997
Job time : 77 secs.
```

```
************************************************************
                       MPSRCH  (TM)
************************************************************

            Release 2.1D John F. Collins, Biocomputing Research Unit
            Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
                   Distribution rights by IntelliGenetics, Inc.

MPsrch_pp   protein - protein database search, using Smith-Waterman algorithm

Run on:            Thu Dec 19 09:27:49 1997;  MasPar time 3.19 Seconds
                                              254.654 Million cell updates/sec
Tabular output not generated.

Title:             >US-08-826-361A-3
Description:       (1-66) from US08826361A.pep
Perfect Score:     535
Sequence:          1 CAVCSDYASGYHYGVWSCEG . KNRRKSCQACRLRKCYEVGM 66

Scoring table:     PAM 150
                   Gap 11

Searched:          101610 seqs, 12294212 residues

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database:          a-geneseq28
                     1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
                     8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
                     14:part14 15:part15 16:part16 17:part17 18:part18
                     19:part19 20:part20 21:part21

Statistics:        Mean 26.304;  Variance 96.939;  scale 0.271

       Pred. No. is the number of results predicted by chance to have a
       score greater than or equal to the score of the result being printed,
       and is derived by analysis of the total score distribution.

                               SUMMARIES
                      %
Result              Query
  No.    Score   Match Length DB  ID            Description            Pred. No.
------------------------------------------------------------------------------
   1      535   100.0    485 21  W14724        Human oestrogen recep  1.89e-48
   2      532    99.4    485 21  W14725        Mouse oestrogen recep  4.13e-48
   3      532    99.4    485 21  W14723        Rat oestrogen recepto  4.13e-48
   4      527    98.5     68 13  R71566        Partial human oestroge 1.52e-47
   5      527    98.5     72  1  P91994        Part of oestrogen rec  1.52e-47
   6      527    98.5     77 10  R53289        Human oestrogen recep  1.52e-47
   7      527    98.5    595  2  P70543        Human oestrogen recep  1.52e-47
   8      500    93.5    334  2  P80925        Sequence of the human  1.67e-44
   9      387    72.3    433  2  P80931        Sequence of human est  7.29e-32
  10      374    69.9    521  2  P80930        Sequence of human est  2.02e-30
  11      363    67.9     66  2  R08250        GTG3A chimeric glucoc  2.23e-29
  12      362    67.7     66  2  R07395        GTG2 mutant glucocort  4.29e-29
  13      356    66.5     66  2  R07396        GTG1 mutant glucocort  1.97e-28
  14      352    65.8     72  1  P91993        Part of progesteron r  5.46e-28
  15      351    65.6     72  1  P91992        Part of mineralocorti  7.04e-28
  16      351    65.6    902  1  P93110        Rat androgen receptor  7.04e-28
  17      351    65.6    902  1  P91006        Rat androgen receptor  7.04e-28
  18      351    65.6    902  3  R12224        Rat androgen receptor  7.04e-28
  19      351    65.6    918  3  R12223        Human androgen recept  7.04e-28
  20      351    65.6    919 21  W14783        Androgen receptor.     7.04e-28
```

```
  21      351    65.6    919  1  P90996        Human androgen recept  7.04e-28
  22      351    65.6    919  1  P93109        Human androgen recept  7.04e-28
  23      349    65.2     66  1  R05139        Human glucocorticoid   1.17e-27
  24      349    65.2     72  1  P91991        Part of glucocorticoi  1.17e-27
  25      349    65.2     93 18  R90714        DNA-binding domain of  1.17e-27
  26      349    65.2    777  2  P80919        Sequence of the prima  1.17e-27
  27      342    63.9    633  3  R12237        TrpE/androgen recepto  6.91e-27
  28      341    63.7    984  2  P80927        Sequence of the human  8.91e-27
  29      338    63.2    447  4  R20466        Mouse retinoic acid r  1.91e-26
  30      338    63.2    454 16  R84723        Human retinoic acid r  1.91e-26
  31      338    63.2    454  2  R10182        Recombinant human gam  1.91e-26
  32      338    63.2    454  2  R10550        Human Retinoic Acid R  1.91e-26
  33      338    63.2    458  4  R20465        Mouse retinoic acid r  1.91e-26
  34      338    63.2    458  2  R10549        Skin-specific murine   1.91e-26
  35      337    63.0    151 10  R55161        Fragment of retinoic   2.45e-26
  36      337    63.0    448 10  R55129        Retinoic acid recepto  2.45e-26
  37      337    63.0    448 12  R66764        Human RAR-beta.        2.45e-26
  38      337    63.0    448 17  R85315        Human hepatoma retino  2.45e-26
  39      333    62.2    151 17  R85316        Human retinoic acid r  6.76e-26
  40      333    62.2    403 17  R89024        RAR-alpha-403 dominan  6.76e-26
  41      333    62.2    448 16  R84725        Murine retinoic acid   6.76e-26
  42      333    62.2    462  5  R27534        RAR-alpha protein.     6.76e-26
  43      333    62.2    462 13  R68023        RAR-alpha.             6.76e-26
  44      333    62.2    462  1  P90295        Retinoic acid recepto  6.76e-26
  45      333    62.2    737  5  R27537        PML/RAR-alpha fusion.  6.76e-26

                              ALIGNMENTS

RESULT    1
ID   W14724 standard; Protein; 485 AA.
AC   W14724;
DT   08-JUN-1997  (first entry)
DE   Human oestrogen receptor beta (ER-beta).
KW   Orphan receptor; oestrogen receptor beta; ER-beta;
KW   nuclear receptor; prostate cancer; benign prostatic hyperplasia;
KW   ovary cancer; cardiovascular disease; osteoporosis;
KW   environment; pollutant.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   Domain          104..169
FT   /label= DNA-binding_domain
FT   Domain          260..457
FT   /label= ligand_binding_domain
PN   WO9709348-A2.
PD   13-MAR-1997.
PF   09-SEP-1996; E03933.
PR   08-SEP-1995; GB-018272.
PR   15-MAR-1996; GB-005550.
PR   11-APR-1996; GB-007532.
PR   08-MAY-1996; GB-009576.
PA   (KARO-) KARO BIO AB.
PI   Enmark E,  Gustafsson J,  Kuiper GG;
DR   WPI; 97-192842/17.
DR   N-PSDB; T62843.
PT   New isolated oestrogen receptor beta - used to develop prods. for
PT   treating e.g. cancers, CNS diseases, osteoporosis or cardiovascular
PT   disease
PS   Claim 1; Fig 13A; 45pp; English.
CC   A novel human orphan receptor (W14724) is related to the known
CC   oestrogen receptor ER-alpha, and has been designated ER-beta.
CC   It is an oestrogen receptor-related nuclear receptor.  The ER-beta
CC   amino acid sequence was deduced from a cDNA clone (T62843) isolated
CC   from a human testis cDNA library.  Rat, human and mouse ER-beta
CC   (W14723-25) can be used to isolate molecules for use in the
CC   treatment of cardiovascular diseases, central nervous system
CC   diseases, osteoporosis, prostate or ovarian cancer or benign
CC   prostatic hyperplasia and to test environmental chemicals for
CC   oestrogenic activity.
SQ   Sequence   485 AA;

  Query Match             100.0%;  Score 535;  DB 21;  Length 485;
  Best Local Similarity 100.0%;  Pred. No. 1.89e-48;
  Matches     66;  Conservative     0;  Mismatches    0;  Indels    0;  Gaps    0;

Db     104 cavcsdyasgyhygvwscegckaffkrsiqghndyicpatnqctidknrrkscqacrlrk 163
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 60

Db     164 cyevgm 169
           ||||||
Qy      61 CYEVGM 66


RESULT    2
ID   W14725 standard; Protein; 485 AA.
AC   W14725;
DT   08-JUN-1997  (first entry)
DE   Mouse oestrogen receptor beta (ER beta).
KW   Orphan receptor; oestrogen receptor beta; ER-beta;
KW   nuclear receptor; prostate cancer; benign prostatic hyperplasia;
KW   ovary cancer; cardiovascular disease; osteoporosis;
KW   environment; pollutant.
OS   Mus sp.
FH   Key              Location/Qualifiers
FT   Domain           104..169
FT   /label= DNA-binding_domain
FT   Domain           260..457
FT   /label= Ligand-binding_domain
PN   WO9709348-A2.
PD   13-MAR-1997.
PF   09-SEP-1996; E03933.
PR   08-SEP-1995; GB-018272.
PR   15-MAR-1996; GB-005550.
PR   11-APR-1996; GB-007532.
PR   08-MAY-1996; GB-009576.
PA   (KARO-) KARO BIO AB.
PI   Enmark E,  Gustafsson J,  Kuiper GG;
DR   WPI; 97-192842/17.
DR   N-PSDB; T62844.
PT   New isolated oestrogen receptor beta - used to develop prods  for
PT   treating e.g. cancers, CNS diseases, osteoporosis or cardiovascular
PT   disease
PS   Claim 1; Fig 14A; 45pp; English.
CC   A novel mouse orphan receptor (W14725) is related to the known
CC   oestrogen receptor ER-alpha, and has been designated ER-beta.
CC   It is an oestrogen receptor-related nuclear receptor.  The ER-beta
CC   amino acid sequence was deduced from a cDNA clone (T62844)
CC   Rat, human and mouse ER-beta (W14723-25) can be used to isolate
CC   molecules for use in the treatment of cardiovascular diseases,
CC   central nervous system diseases, osteoporosis, prostate or ovarian
CC   cancer or benign prostatic hyperplasia and to test environmental
CC   chemicals for oestrogenic activity.
SQ   Sequence   485 AA;

  Query Match             99.4%;  Score 532;  DB 21;  Length 485;
  Best Local Similarity   98.5%;  Pred. No. 4.13e-48;
  Matches     65;  Conservative     1;  Mismatches    0;  Indels    0;  Gaps    0;

Db     104 cavcsdyasgyhygvwscegckaffkrsiqghndyicpatnqctidknrrkncqacrlrk 163
           ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Qy       1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 60

Db     164 cyevgm 169
           ||||||
Qy      61 CYEVGM 66


RESULT    3
ID   W14723 standard; Protein; 485 AA.
AC   W14723;
DT   08-JUN-1997  (first entry)
DE   Rat oestrogen receptor beta (ER-beta).
KW   Orphan receptor; oestrogen receptor beta; ER-beta;
KW   nuclear receptor; prostate cancer; benign prostatic hyperplasia;
KW   ovary cancer; cardiovascular disease; osteoporosis;
KW   environment; pollutant.
OS   Rattus sp.
FH   Key              Location/Qualifiers
FT   Domain           104..169
FT   /label= DNA-binding_domain
FT   Domain           259..457
FT   /label= Ligand-binding_domain
PN   WO9709348-A2.
PD   13-MAR-1997.
PF   09-SEP-1996; E03933.
PR   08-SEP-1995; GB-018272.
PR   15-MAR-1996; GB-005550.
PR   11-APR-1996; GB-007532.
PR   08-MAY-1996; GB-009576.
PA   (KARO-) KARO BIO AB.
PI   Enmark E,  Gustafsson J,  Kuiper GG;
DR   WPI; 97-192842/17.
DR   N-PSDB; T62842.
PT   New isolated oestrogen receptor beta - used to develop prods. for
PT   treating e.g. cancers, CNS diseases, osteoporosis or cardiovascular
PT   disease
PS   Claim 1; Fig 1; 45pp; English
CC   A novel rat orphan receptor (W14723) is related to the known
CC   oestrogen receptor ER-alpha, and has been designated ER-beta.
CC   It is an oestrogen receptor-related nuclear receptor.  The ER-beta
CC   amino acid sequence was deduced from a cDNA clone (T62842) isolated
CC   from a rat prostate cDNA library.  Rat, human and mouse ER-beta
CC   (W14723-25) can be used to isolate molecules for use in the
CC   treatment of cardiovascular diseases, central nervous system
CC   diseases, osteoporosis, prostate or ovarian cancer or benign
CC   prostatic hyperplasia and to test environmental chemicals for
CC   oestrogenic activity.
SQ   Sequence   485 AA;

  Query Match             99.4%;  Score 532;  DB 21;  Length 485;
  Best Local Similarity   98.5%;  Pred. No. 4.13e-48;
  Matches     65;  Conservative     1;  Mismatches    0;  Indels    0;  Gaps    0;

Db     104 cavcsdyasgyhygvwscegckaffkrsiqghndyicpatnqctidknrrkncqacrlrk 163
           ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Qy       1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 60

Db     164 cyevgm 169
           ||||||
Qy      61 CYEVGM 66


RESULT    4
ID   R71566 standard; Protein; 68 AA.
AC   R71566;
DT   31-OCT-1995  (first entry)
DE   Partial human estrogen receptor.
KW   nuclear receptor protein; osteosarcoma; EAR-1r; generation; growth;
KW   differentiation; human; erb-A related gene; estrogen receptor.
OS   Homo sapiens.
PN   J07033798-A.
PD   03-FEB-1995.
PF   20-JUL-1993; 200333.
PR   20-JUL-1993; JP-200333.
PA   (MITU ) MITSUBISHI KASEI CORP.
DR   WPI; 95-110650/15.
DR   N-PSDB; Q85976.
PT   Nuclear receptor protein derived from osteo:blastic osteo:sarcoma
PT   - encoded by the EAR-1r gene, for investigation of generation of
PT   fat soluble active substances
PS   Disclosure; Page 6; 8pp; Japanese.
CC   This partial estrogen receptor (a member of a receptor superfamily)
CC   amino acid sequence (positions 185-252) is encoded by the cDNA shown
CC   in Q85976. It is disclosed in the patent specification. The
CC   invention concerns an EAR-1r gene (human erb-A related gene)
CC   encoding a nuclear receptor protein (R71565) derived from human
CC   osteosarcoma (HOS) cells. The expression of EAR-1r is most abundant
CC   in the brain and was detectedd slightly in heart and placenta. The
CC   receptor protein may be used for the investigation of generation,
CC   differentiation and growth of fat soluble physiological substances
SQ   Sequence   68 AA;

  Query Match             98.5%;  Score 527;  DB 13;  Length 68;
  Best Local Similarity  97.0%;  Pred  No  1.52e-47;
  Matches    64;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Db         1 cavcndyasgyhygvwscegckaffkrsiqghndymcpatnqctidknrrkscqacrlrk 60
             ||||:||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Qy         1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 60

Db        61 cyevgm 66
             ||||||
Qy        61 CYEVGM 66


RESULT    5
ID   P91994 standard; protein; 72 BP.
AC   P91994;
DT   05-MAR-1990  (first entry)
DE   Part of oestrogen receptor.
KW   Oestrogen receptor; cystein-rich DNA binding domain;
OS   Homo sapiens.
PN   WO8909223-A.
PD   05-OCT-1989.
PF   24-MAR-1989; U01238
PR   30-MAR-1988; US-176107.
PA   (ARCH-) Arch Development Corp.
PI   Liao S, Chang C;
DR   WPI; 89-309501/42.
PT   New DNA encoding new androgen receptor and TR2 polypeptide(s) - able
PT   to bind DNA, and derived antibodies, useful for receptor assay and
PT   purification.
PS   Disclosure; fig. 5; 60pp; English
CC   The sequence is part of the oestrogen receptor and is homologous
CC   with the cystein-rich DNA binding domain of human androgen receptor,
CC   glucocorticoid receptor, mineralocorticoid receptor, progesterone
CC   receptor, TR2, rat AR, chick vitamin D receptor and the v-erb oncogene
CC   product of avian erythroblastosis virus.
CC   See also P91991 - P91995.
SQ   Sequence   72 AA;

  Query Match             98.5%;  Score 527;  DB 1;  Length 72;
  Best Local Similarity  97.0%;  Pred  No  1.52e-47;
  Matches    64;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Db         3 cavcndyasgyhygvwscegckaffkrsiqghndymcpatnqctidknrrkscqacrlrk 62
             ||||:||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Qy         1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 60

Db        63 cyevgm 68
             ||||||
Qy        61 CYEVGM 66


RESULT    6
ID   R53289 standard; protein; 77 AA.
AC   R53289;
DT   30-DEC-1994  (first entry)
DE   Human oestrogen receptor DNA-binding domain (185-261).
KW   Human; oestrogen receptor; DNA-binding domain; hER;
KW   monoclonal antibody; diagnosis; antiserum; cancer; hapten;
KW   immunogen.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   Peptide         17..31
FT   /label= peptide_1
FT   /note= "hER 201-215"
FT   Peptide         46..61
FT   /label= peptide_2
FT   /note= "hER 231-245"
FT   Peptide         63..77
FT   /label= peptide_3
FT   /note= "hER 247-261"
PN   US5312752-A.
PD   17-MAY-1994.
PF   31-JUL-1989; 388091.
PR   31-JUL-1989; US-388091
PR   01-NOV-1991; US-784626.
PA   (UYBO-) UNIV BOSTON.
PI   Iraish AM,  Wotiz HH;
DR   WPI; 94-159123/19.
PT   Monoclonal antibody to human oestrogen receptor protein -
PT   specific for activated forms, useful for diagnosis, esp. in
PT   breast cancer assessment
PS   Disclosure; Fig 1; 41pp; English.
CC   The DNA-binding region of human oestrogen receptor (hER) is
CC   encompassed by the amino acids 180-251 in the native protein.
CC   Three peptides chosen as haptens and immunogens to obtain specific
CC   antisera, are indicated in the features table.
SQ   Sequence   77 AA;

  Query Match             98.5%;  Score 527;  DB 10;  Length 77;
  Best Local Similarity  97.0%;  Pred. No. 1.52e-47;
  Matches    64;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Db         1 cavcndyasgyhygvwscegckaffkrsiqghndymcpatnqctidknrrkscqacrlrk 60
             ||||:||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Qy         1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 60

Db        61 cyevgm 66
             ||||||
Qy        61 CYEVGM 66


RESULT    7
ID   P70543 standard, protein; 595 AA.
AC   P70543;
DT   24-FEB-1991  (first entry)
DE   Human oestrogen receptor protein.
KW   Human oestrogen receptor protein, expression system; assay;
OS   Homo sapiens.
PN   WO8705049-A.
PD   27-AUG-1987.
PF   18-FEB-1987; U00341.
PR   20-FEB-1986; US-833829.
PA   (CALB-) California Biotechnology Inc.
PI   Shine J;
DR   WPI; 87-250213/35.
DR   N-PSDB; N70880.
PT   Expression system for vertebrate steroid receptor protein - comprising
PT   DNA sequence encoding the protein linked to control sequences in
PT   eucaryotic hosts.
PS   Disclosure; Fig. 1-1 - 1-2; 26pp; English.
CC   steroid receptor protein;
CC   The protein is expressed in an expression system from cDNA operably
CC   linked to control sequences compatible with eukaryotic host cells.
CC   This method allows expression under conditions which favour appropriate
CC   post-translational processing.  It also produces large amts. of purified
CC   protein useful in the design of agonist and antagonist cpds. for study
CC   of the mechanism of action of the steroid binding proteins in general,
CC   and for use in diagnostic assays for the proteins or antibodies to
CC   them.  These assays are important in, eg the diagnosis of tumour
CC   sensitivities to steroid metabolism.  Suitable host cells are VERO,
CC   HeLa and CHO cells.
SQ   Sequence   595 AA;

  Query Match             98.5%;  Score 527;  DB 2;  Length 595;
  Best Local Similarity  97.0%;  Pred. No. 1.52e-47;
  Matches    64;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Db       185 cavcndyasgyhygvwscegckaffkrsiqghndymcpatnqctidknrrkscqacrlrk 244
             ||||:||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Qy         1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 60
```

```
Db     245 cyrvgm 250
           ||||||
Qy      61 CYEVGM 66


RESULT    8
ID   P80925 standard; protein; 334 AA.
AC   P80925;
DT   29-DEC-1990  (first entry)
DE   Sequence of the human oestrogen receptor
KW   Hormone receptor; hormone-binding; transcription activation;.
OS   Homo sapiens.
PN   WO8803168-A.
PD   05-MAY-1988.
PF   23-OCT-1987; U02782.
PR   20-OCT-1987; US-108471.
PA   (SALK) Salk Inst for Biol Stud.
PI   Evans RM, Weinberger CA, Hollenberg SM, Giguere V;
DR   WPI; 88-133242/19.
PT   Recombinant DNA encoding hormone recpetors -
PT   comprising glucocorticoid  mineralocorticoid, thyroid hormone
PT   and novel hormone receptors
PS   Example; Fig III-2; 243pp; English.
CC   DNAs encoding hormone receptors and the hormone receptors themselves are
CC   claimed. The DNA can be used to make the hormone receptor proteins and
CC   functional modified forms in quantities not previously possible. The
CC   receptor proteins can be used to screen cpds. for receptor-agonist or
CC   receptor-antagonist activity. They can also be used in diagnostic assays.
SQ   Sequence   334 AA;

  Query Match             93.5%;  Score 500;  DB 2;  Length 334;
  Best Local Similarity   93.9%;  Pred  No. 1.67e-44;
  Matches    62;  Conservative    2;  Mismatches    1;  Indels    1;  Gaps    1;

Db       2 cavcndyasgyhygvwscegckaffkrs-qghndymcpatnqctidknrrkscqacrlrk 60
           ||||:||||||||||||||||||||||| ||||||:||||||||||||||||||||||||
Qy       1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 60

Db      61 cyrvgm 66
           || |||
Qy      61 CYEVGM 66


RESULT    9
ID   P80931 standard; protein; 433 AA.
AC   P80931;
DT   29-DEC-1990  (first entry)
DE   Sequence of human estrogen-related recepter protein (hERR2)
KW   Hormone receptor; hormone-binding; transcription activation.
OS   Homo sapiens.
PN   WO8803168-A.
PD   05-MAY-1988.
PF   23-OCT-1987; U02782.
PR   20-OCT-1987; US-108471.
PA   (SALK) Salk Inst for Biol Stud.
PI   Evans RM, Weinberger CA, Hollenberg SM, Giguere V;
DR   WPI; 88-133242/19.
DR   N-PSDB; N80922.
PT   Recombinant DNA encoding hormone recpetors -
PT   comprising glucocorticoid, mineralocorticoid, thyroid hormone
PT   and novel hormone receptors
PS   Claim 17; Fig V-2(B)-1 and -2; 243pp, English.
CC   DNAs encoding hormone receptors and the hormone receptors themselves are
CC   claimed. The DNA can be used to make the hormone receptor proteins and
CC   functional modified forms in quantities not previously possible. The
CC   receptor proteins can be used to screen cpds. for receptor-agonist or
CC   receptor-antagonist activity. They can also be used in diagnostic assays.
CC   Also claimed is pure DNA encoding oestrogen-related receptors hERR1 and
CC   hERR2 and hERR1 and hERR2. The new hERR1 and hERR2 receptors will provide
CC   the basis for development of an assay system that will lead to the
CC   identification of novel hormones.
SQ   Sequence   433 AA;

  Query Match             72.3%;  Score 387;  DB 2;  Length 433;
  Best Local Similarity   72.7%;  Pred. No. 7.29e-32;
  Matches    48;  Conservative    7;  Mismatches   11;  Indels    0;  Gaps    0;

Db     103 clvcgdiasgyhygvasceackaffkrtiqgnieyscpatneceitkrrkscqacrfmk 162
           | ||:| |||||||| |||:||||||||:|||: :| |||||:| | | |||||||||: |
Qy       1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 60

Db     163 clkvgm 168
           |  |||
Qy      61 CYEVGM 66


RESULT   10
ID   P80930 standard; protein; 521 AA.
AC   P80930;
DT   10-MAR-1993  (revised)
DT   29-DEC-1990  (first entry)
DE   Sequence of human estrogen-related receptor protein (hERR1)
KW   Hormone receptor; hormone binding; transcription activation.
OS   Homo sapiens.
PN   WO8803168-A.
PD   05-MAY-1988.
PF   23-OCT-1987; U02782.
PR   20-OCT-1987; US-108471.
PA   (SALK) Salk Inst for Biol Stud.
PI   Evans RM, Weinberger CA, Hollenberg SM, Giguere V;
DR   WPI; 88-133242/19.
DR   N-PSDB; N80921.
PT   Recombinant DNA encoding hormone recpetors -
PT   comprising glucocorticoid, mineralocorticoid, thyroid hormone
PT   and novel hormone receptors
PS   Claim 17; Fig V-1(B)-1 and -2; 243pp, English.
CC   DNAs encoding hormone receptors and the hormone receptors themselves are
CC   claimed. The DNA can be used to make the hormone receptor proteins and
CC   functional modified forms in quantities not previously possible. The
CC   receptor proteins can be used to screen cpds. for receptor-agonist or
CC   receptor-antagonist activity. They can also be used in diagnostic assays.
CC   Also claimed is pure DNA encoding oestrogen-related receptors hERR1 and
CC   hERR2 and hERR1 and hERR2. The new hERR1 and hERR2 receptors will provide
CC   the basis for development of an assay system that will lead to the
CC   identification of novel hormones.
SQ   Sequence   521 AA;

  Query Match             69.9%;  Score 374;  DB 2;  Length 521;
  Best Local Similarity   69.7%;  Pred. No. 2.02e-30;
  Matches    46;  Conservative    8;  Mismatches   12;  Indels    0;  Gaps    0;

Db     176 clvcgdvasgyhygvasceackaffkrtiqgsieyscpasneceitkrrrkacqacrftk 235
           | ||:| |||||||| |||:||||||||:|||   :| |||:|:| | | |||:|||||: |
Qy       1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 60

Db     236 clrvgm 241
           |  |||
Qy      61 CYEVGM 66


RESULT   11
ID   R08250 standard, protein; 66 AA.
AC   R08250,
DT   01-FEB-1991  (first entry)
DE   GTG3A chimeric glucocorticoid/thyroid receptor.
KW   Steroid; thyroid; hormone response element; HER;
KW   ligand identification; chimera; orphan receptor protein, ORP.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   Domain          1..21
FT   /label=finger_1
FT   Region          22..36
FT   /label=linker
FT   Domain          39..61
FT   /label=finger_2
PN   WO9011273-A.
```

```
PD   04-OCT-1990.
PF   16-MAR-1990, U01428.
PR   17-MAR-1989; US-325240.
PA   (SALK ) SALK INST FOR BIOL STUD.
PI   Evans RM, Kazuhiko U;
DR   WPI; 90-320209/42
PT   Steroid and thyroid hormone response elements - used to develop
PT   assays for identifying ligands for orphan hormone receptors
PS   Disclosure; Fig 2; 48pp; English.
CC   To identify amino acids that allow the receptors to discriminate
CC   between their cognate HREs, a variety of chimeric DNA binding
CC   domains were constructed.
CC   A first proximal (P) region follows C3 and includes the three amino
CC   acid cluster GS--V in the GR and EG--G in the TR.  These are part
CC   of a ten amino acid segment that is highly conserved and includes
CC   6 invariant amino acids.  When these three amino acids are
CC   switched, the resultant chimera (GTG3A) is nonfunctional on either
CC   MTV-LUC or deltaMTV-PREp-LUC.
CC   The second or distal (D) region of importance is a cluster of 5
CC   amino acids between C5 and C6: AGRND in hGR and KYEGK in hTRbeta.
CC   See also R08245-51, R07366-86, R07391, -95 and -96.
SQ   Sequence   66 AA;

  Query Match             67.9%;  Score 363;  DB 2;  Length 66;
  Best Local Similarity   62.1%;  Pred. No. 3.32e-29;
  Matches   41;  Conservative   14;  Mismatches   11;  Indels    0;  Gaps    0;

Db       1 clvcsdeasgchygvltcegckgffkravegqhnylcagrndciidkirrkncpacryrk 60
           | |||| ||| |||| :|||||:||||:..|...|.|:: |:| ||| |||:| ||| ||
Qy       1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 60

Db      61 clqagm 66
           | : ||
Qy      61 CYEVGM 66


RESULT   12
ID   R07395 standard; protein; 66 AA.
AC   R07395;
DT   01-FEB-1991  (first entry)
DE   GTG2 mutant glucocorticoid/thyroid receptor.
KW   Steroid; thyroid; hormone response element; HER;
KW   ligand identification; mutant; orphan receptor protein; ORP.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   Domain          1..21
FT   /label=finger_1
FT   Region          22..36
FT   /label=linker
FT   Domain          39..61
FT   /label=finger_2
PN   WO9011273-A.
PD   04-OCT-1990.
PF   16-MAR-1990, U01428.
PR   17-MAR-1989, US 325240.
PA   (SALK ) SALK INST FOR BIOL STUD.
PI   Evans RM, Kazuhiko U;
DR   WPI; 90-320209/42
PT   Steroid and thyroid hormone response elements - used to develop
PT   assays for identifying ligands for orphan hormone receptors
PS   Disclosure; Fig 3; 48pp; English.
CC   Although the triple mutant GTG3A (R08250) selectively recognises
CC   the ERE, the double mutant GTG2 is active upon both MTV and ERE
CC   reporters.  This induction is dependent on the HRE since this
CC   receptor fails to activate either deltaMTV or TREp promoter.
CC   GTG2 lacks the T3Rbeta D element.
CC   See also R08245-51, R07366-86, R07391, -95 and -96.
SQ   Sequence   66 AA;

  Query Match             67.7%;  Score 362;  DB 2;  Length 66;
  Best Local Similarity   62.1%;  Pred. No. 4.29e-29;
  Matches   41;  Conservative   13;  Mismatches   12;  Indels    0;  Gaps    0;

Db       1 clvcsdeasgchygvltcegckvffkravegqhnylcagrndciidkirrkncpacryrk 60
           | |||| ||| |||| :||||| ||||:::|:::|:|:: |:| ||| |||:| ||| ||
Qy       1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 60

Db      61 clqagm 66
           | : ||
Qy      61 CYEVGM 66


RESULT   13
ID   R07396 standard; protein; 66 AA
AC   R07396;
DT   01-FEB-1991  (first entry)
DE   GTG1 mutant glucocorticoid/thyroid receptor.
KW   Steroid; thyroid; hormone response element; HER;
KW   ligand identification, mutant; orphan receptor protein; ORP.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   Domain          1..21
FT   /label=finger_1
FT   Region          22..36
FT   /label=linker
FT   Domain          39..61
FT   /label=finger_2
PN   WO9011273-A.
PD   04-OCT-1990.
PF   16-MAR-1990, U01428.
PR   17-MAR-1989; US-325240.
PA   (SALK ) SALK INST FOR BIOL STUD
PI   Evans RM, Kazuhiko U;
DR   WPI, 90-320209/42.
PT   Steroid and thyroid hormone response elements - used to develop
PT   assays for identifying ligands for orphan hormone receptors
PS   Disclosure; Fig 3; 48pp; English.
CC   Although the triple mutant GTG3A (R08250) selectively recognises
CC   the ERE, the single mutant GTG1 is active upon both MTV and ERE
CC   reporters.  The amino acid change leaves GRE recognition normal but
CC   fosters clear recognition of the ERE.  As with the GTG2 mutant
CC   (P7395) the induction is dependent on the HRE since this
CC   receptor fails to activate either deltaMTV or TREp promoter.
CC   GTG1 lacks the T3Rbeta D element.
CC   See also R08245-51, R07366-86, R07391, -95 and -96
SQ   Sequence   66 AA;

  Query Match             66.5%;  Score 356;  DB 2;  Length 66;
  Best Local Similarity   60.6%,  Pred. No. 1.97e-28;
  Matches   40,  Conservative   14;  Mismatches   12,  Indels    0;  Gaps    0;

Db       1 clvcsdeasgchygvltcesckvffkravegqhnylcagrndciidkirrkncpacryrk 60
           | |||| ||| |||| :||:|| ||||:::|::|:|:: |:| ||| |||:| ||| ||
Qy       1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRLRK 60

Db      61 clqagm 66
           | : ||
Qy      61 CYEVGM 66


RESULT   14
ID   P91993 standard; protein; 72 BP
AC   P91993;
DT   05-MAR-1990  (first entry)
DE   Part of progesteron receptor.
KW   Progesterone receptor; cystein-rich DNA binding domain;
OS   Homo sapiens.
PN   WO8909223-A.
PD   05-OCT-1989.
PF   24-MAR-1989; U01238.
PR   30-MAR-1988; US-176107
PA   (ARCH-) Arch Development Corp.
PI   Liao S, Chang C;
DR   WPI; 89-309501/42.
PT   New DNA encoding new androgen receptor and TR2 polypeptide(s) - able
PT   to bind DNA, and derived antibodies, useful for receptor assay and
PT   purification.
PS   Disclosure; fig. 5; 60pp; English.
CC   The sequence is part of the progesterone receptor and is homologous
CC   with the cystein-rich DNA binding domain of human androgen receptor,
CC   glucocorticoid receptor, mineralocorticoid receptor, oestrogen receptor,
CC   TR2, rat AR, chick vitamin D receptor and the v-erb oncogene product of
CC   avian erythroblastosis virus.
CC   See also P91991 - P91995.
SQ   Sequence   72 AA;

  Query Match             65.8%;  Score 352;  DB 1;  Length 72;
  Best Local Similarity  56.1%;  Pred. No. 5.46e-28;
  Matches   37;  Conservative   17;  Mismatches  12;  Indels   0;  Gaps   0;

Db        3 clicgdeasgchygvltcgsckvffkramegqhnylcagrndcivdkirrkncpacrlrk 62
            | :|:| ||| |||| :| :|| ||||:::|:::|:|:: |:| :|| |||:| ||||||
Qy        1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 60

Db       63 ceqagm 68
            | : ||
Qy       61 CYEVGM 66


RESULT   15
ID   P91992 standard; protein; 72 BP.
AC   P91992;
DT   05-MAR-1990  (first entry)
DE   Part of mineralocorticoid receptor.
KW   Mineralocorticoid receptor; cystein-rich DNA binding domain;
OS   Homo sapiens.
PN   WO8909223-A.
PD   05-OCT-1989.
PF   24-MAR-1989; U01238.
PR   30-MAR-1988; US-176107.
PA   (ARCH-) Arch Development Corp.
PI   Liao S, Chang C;
DR   WPI; 89-309501/42.
PT   New DNA encoding new androgen receptor and TR2 polypeptide(s) - able
PT   to bind DNA, and derived antibodies, useful for receptor assay and
PT   purification.
PS   Disclosure; fig. 5; 60pp; English.
CC   The sequence is part of the mineralocorticoid receptor and is homologous
CC   with the cystein-rich DNA binding domain of human androgen receptor,
CC   glucocorticoid receptor, progesterone receptor, oestrogen receptor,
CC   TR2, rat AR, chick vitamin D receptor and the v-erb oncogene product of
CC   avian erythroblastosis virus.
CC   See also P91991 - P91995.
SQ   Sequence   72 AA;

  Query Match             65.6%;  Score 351;  DB 1;  Length 72;
  Best Local Similarity  59.1%;  Pred. No. 7.04e-28;
  Matches   39;  Conservative   15;  Mismatches  12;  Indels   0;  Gaps   0;

Db        3 clvcgdeasgchygvvtcgsckvffkravegqhnylcagrndciidkirrkncpacrlrk 62
            | ||:| ||| |||| :| :|| ||||:::|:::|:|:: |:| ||| |||:| ||||||
Qy        1 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 60

Db       63 clqagm 68
            | : ||
Qy       61 CYEVGM 66

Search completed: Thu Dec 18 08:28:02 1997
Job time : 13 secs.
```

```
**************************************************************************

        MPSRCH                                                      (TM)

**************************************************************************

          Release 2 1D John F  Collins, Biocomputing Research Unit.
          Copyright (c) 1993  1994  1995  University of Edinburgh, U.K.
                    Distribution rights by IntelliGenetics, Inc.

MPsrch_pp   protein - protein database search  using Smith-Waterman algorithm

Run on:            Thu Dec 18 09:14.20 1997.  MasPar time 12.50 Seconds
                                              521.324 Million cell updates/sec
Tabular output not generated.

Title:             >US-08-826-361A-25
Description:       (1-530) from US08826361A pep
Perfect Score:     3904
Sequence:          1 MDIKNSPSSLNSPSSYNCSQ          ECSPAEDSKSKEGSQNPQSQ 530

Scoring table:     PAM 150
                   Gap 11

Searched:          101610 seqs, 12294212 residues

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database:          a-geneseq28
                     1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
                     8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
                     14:part14 15:part15 16:part16 17:part17 18:part18
                     19:part19 20:part20 21:part21

Statistics:        Mean 36.091;  Variance 153.377;  scale 0.235

      Prod. No  is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                SUMMARIES
                 %
Result         Query
  No.   Score  Match Length DB  ID            Description            Pred. No
--------------------------------------------------------------------------------
   1    3582   91.8    485 21  W14724        Human oestrogen recep  0.00e+00
   2    3274   83.9    485 21  W14723        Rat oestrogen recepto  0.00e+00
   3    3250   83.2    485 21  W14725        Mouse oestrogen recep  0.00e+00
   4    1671   42.8    595  2  P70543        Human oestrogen recep  3.53e-151
   5    1424   36.5    334  2  P80925        Sequence of the human  1.87e-126
   6     807   20.7    433  2  P80931        Sequence of human est  3.05e-65
   7     800   20.5    521  2  P80930        Sequence of human est  1.49e-64
   8     614   15.7    777  2  P80919        Sequence of the prima  2.11e-46
   9     576   14.8    902  3  R12224        Rat androgen receptor  1.00e-42
  10     570   14.6    902  1  P93110        Rat androgen recepto   2.81e-42
  11     570   14.6    902  1  P91006        Rat androgen receptor  2.81e-42
  12     570   14.6    918  2  R12223        Human androgen recept  2.81e-42
  13     567   14.5    919  1  P90996        Human androgen recept  7.42e-42
  14     567   14.5    919 21  W14782        Androgen receptor      7.42e-42
  15     566   14.5    919  1  P93109        Human androgen recept  9.27e-42
  16     563   14.4     77 10  R53289        Human oestrogen recep  1.81e-41
  17     546   14.0     72  1  P91994        Part of oestrogen rec  7.85e-40
  18     545   14.0    920  3  R12230        TrpE/androgen recepti  9.80e-40
  19     536   13.7     69 13  R71566        Partial human estroge  7.22e-39
  20     533   13.7    533 13  R72483        Human H-2RIIBP.        1.40e-38
```

```
  21     533   13.7    533  8  R39466        hRXR beta1.            1.40e-38
  22     529   13.6    446  2  R11837        H2RIIBP nuclear hormo  3.39e-38
  23     529   13.6    446 13  R72484        H 2RIIBP.              3.39e-38
  24     528   13.5    533  8  R39469        hRXR-beta2.            4.23e-38
  25     525   13.4    448  9  R39467        mRXR-beta.             8.21e-38
  26     521   13.3    984  2  P80927        Sequence of the human  1.99e-37
  27     514   13.2    467  8  R39470        mRXR-alpha.            9.34e-37
  28     514   13.2    467  3  R13799        Mouse RXR-alpha retin  9.34e-37
  29     509   13.0    462  2  R13798        Human RXR-alpha retin  2.82e-36
  30     499   12.8    463  9  R39471        mRXR-gamma.            2.31e-34
  31     486   12.4    463  3  R13800        Mouse RXR-gamma retin  4.48e-34
  32     473   12.1    463 20  W00605        Human retinoid X rece  7.80e-33
  33     471   12.1    599  9  R49631        Sequence of nuclear r  1.21e-32
  34     449   11.5    447  4  R20466        Mouse retinoic acid r  1.50e-30
  35     450   11.5    458  2  R10549        Skin-specific murine   1.21e-30
  36     449   11.5    459  4  R20465        Mouse retinoic acid r  1.50e-30
  37     444   11.4    448  2  R10548        Murine Retinoic Acid   4.49e-30
  38     444   11.4    449 16  R84725        Murine retinoic acid   4.49e-30
  39     442   11.3    449 10  R55129        Retinoic acid recepto  6.95e-30
  40     441   11.3    448 17  R85315        Human hepatoma retino  8.64e-30
  41     441   11.3    448 12  R66764        Human RAR-beta.        8.64e-30
  42     441   11.3    454  2  R10192        Recombinant human gam  8.64e-30
  43     441   11.3    454  2  R10550        Human Retinoic Acid R  8.64e-30
  44     441   11.3    454 16  R84723        Human retinoic acid r  8.64e-30
  45     443   11.3    458 16  R84726        Murine retinoic acid   7.58e-30

                                ALIGNMENTS

RESULT     1
ID   W14724 standard; Protein; 485 AA.
AC   W14724;
DT   08-JUN-1997  (first entry)
DE   Human oestrogen receptor beta (ER-beta).
KW   Orphan receptor; oestrogen receptor beta; ER-beta;
KW   nuclear receptor; prostate cancer; benign prostatic hyperplasia;
KW   ovary cancer; cardiovascular disease; osteoporosis;
KW   environment; pollutant.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   Domain          104..169
FT   /label= DNA-binding_domain
FT   Domain          260..457
FT   /label= Ligand-binding_domain
PN   WO9709348-A2.
PD   13-MAR-1997.
PF   09-SEP-1996; E03933.
PR   08-SEP-1995; GB-018272.
PR   15-MAR-1996; GB-005550.
PR   11-APR-1996; GB-007532.
PR   08-MAY-1996; GB-009576.
PA   (KARO-) KARO BIO AB.
PI   Enmark E,  Gustafsson J,  Kuiper GG,
DR   WPI; 97-192842/17.
DR   N-PSDB; T62843.
PT   New isolated oestrogen receptor beta - used to develop prods. for
PT   treating e.g. cancers, CNS diseases, osteoporosis or cardiovascular
PT   disease
PS   Claim 1; Fig 13A; 45pp; English.
CC   A novel human orphan receptor (W14724) is related to the known
CC   oestrogen receptor ER-alpha, and has been designated ER-beta.
CC   It is an oestrogen receptor-related nuclear receptor.  The ER-beta
CC   amino acid sequence was deduced from a cDNA clone (T62843) isolated
CC   from a human testis cDNA library.  Rat, human and mouse ER-beta
CC   (W14723-25) can be used to isolate molecules for use in the
CC   treatment of cardiovascular diseases, central nervous system
CC   diseases, osteoporosis, prostate or ovarian cancer or benign
CC   prostatic hyperplasia and to test environmental chemicals for
CC   oestrogenic activity.
SQ   Sequence   485 AA;

  Query Match              91.8%,  Score 3582,  DB 21;  Length 485;
  Best Local Similarity   99.6%,  Pred. No. 0.00e+00;
  Matches   484;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Db        1 mtfyspavmnysipsnvtnleggpgrqttspnvlwptpghlsplvvhrqlshlyaepqks 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       46 MTFYSPAVMNYSIPSNVTNLEGGPGRQTTSPNVLWPTPGHLSPLVVHRQLSHLYAEPQKS 105

Db       61 pwcearslehtlpvnretlkrkvsgnrcaspvtgpgskrdahfcavcsdyasgyhygvws 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      106 PWCEARSLEHTLPVNRETLKRKVSGNRCASPVTGPGSKRDAHFCAVCSDYASGYHYGVWS 165

Db      121 cegckaffkrsiqghndyicpatnqctidknrrkscqacrlrkcyevgmvkcgsrrercg 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      166 CEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCG 225

Db      181 yrlvrrqrsadeqlhcagkakrsgghaprvrelllldalspeqlvltlleaeppphvlisrp 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      226 YRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQLVLTLLEAEPPHVLISRP 285

Db      241 sapfteasmmmsltkladkelvhmiswakkipgfvelslfdqvrllescwmevlmmglmw 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      286 SAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMW 345

Db      301 rsidhpgklifapdlvldrdegkcvegileifdmllattsrfrelklqhkeylcvkamil 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      346 RSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMIL 405

Db      361 lnssmyplvtatqdadssrklahllnavtdalvwviaksgissqqqsmrlanllmllshv 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      406 LNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHV 465

Db      421 rhasnkgmehllnmkcknvvpvydlllemlnahvlrgckssitgsecspaedskskegsq 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      466 RHASNKGMEHLLNMKCKNVVPVYDLLLEMLNAHVLRGCKSSITGSECSPAEDSKSKEGSQ 525

Db      481 nlqsq 485
            | |||
Qy      526 NPQSQ 530

RESULT    2
ID   W14723 standard; Protein; 485 AA.
AC   W14723;
DT   08-JUN-1997  (first entry)
DE   Rat oestrogen receptor beta (ER-beta).
KW   Orphan receptor; oestrogen receptor beta; ER-beta;
KW   nuclear receptor; prostate cancer; benign prostatic hyperplasia;
KW   ovary cancer; cardiovascular disease; osteoporosis;
KW   environment; pollutant.
OS   Rattus sp.
FH   Key             Location/Qualifiers
FT   Domain          104..169
FT   /label= DNA-binding_domain
FT   Domain          259..457
FT   /label= Ligand-binding_domain
PN   WO9709348-A2.
PD   13-MAR-1997.
PF   09-SEP-1996; E03933.
PR   08-SEP-1995; GB-018272.
PR   15-MAR-1996; GB-005550.
PR   11-APR-1996; GB-007532.
PR   08-MAY-1996; GB-009576.
PA   (KARO-) KARO BIO AB.
PI   Enmark E,  Gustafsson J,  Kuiper GG;
DR   WPI; 97-192842/17.
DR   N-PSDB; T62842.
PT   New isolated oestrogen receptor beta - used to develop prods. for
PT   treating e.g. cancers, CNS diseases, osteoporosis or cardiovascular
PT   disease
PS   Claim 1; Fig 1; 45pp; English.
CC   A novel rat orphan receptor (W14723) is related to the known
CC   oestrogen receptor ER-alpha, and has been designated ER-beta.
CC   It is an oestrogen receptor-related nuclear receptor.  The ER-beta
CC   amino acid sequence was deduced from a cDNA clone (T62842) isolated
CC   from a rat prostate cDNA library.  Rat, human and mouse ER-beta
CC   (W14723-25) can be used to isolate molecules for use in the
CC   treatment of cardiovascular diseases, central nervous system
CC   diseases, osteoporosis, prostate or ovarian cancer or benign
CC   prostatic hyperplasia and to test environmental chemicals for
CC   oestrogenic activity.
SQ   Sequence   485 AA;

  Query Match             83.9%;  Score 3274;  DB 21;  Length 485;
  Best Local Similarity   88.7%;  Pred. No. 0.00e+00;
  Matches   430;  Conservative   35;  Mismatches   20;  Indels    0;  Gaps    0;

Db        1 mtfyspavmnysvpgstsnleggpvrlstspnvlwptsghlsplathcqssllyaepqks 60
            ||||||||||||:|:: :||:|||  |  :||||||||||:||||||     |  |  |||||||
Qy       46 MTFYSPAVMNYSIPSNVTNLEGGPGRQTTSPNVLWPTPGHLSPLVVHRQLSHLYAEPQKS 105

Db       61 pwcearslehtlpvnretlkrklsgsscaspvtspnakrdahfcpvcsdyasgyhygvws 120
            |||||||||||||||||||||||||:||: ||||||:|  :||||||||:||||||||||||||
Qy      106 PWCEARSLEHTLPVNRETLKRKVSGNRCASPVTGPGSKRDAHFCAVCSDYASGYHYGVWS 165

Db      121 cegckaffkrsiqghndyicpatnqctidknrrkscqacrlrkcyevgmvkcgsrrercg 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      166 CEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCG 225

Db      181 yrivrrqrssseqvhclskakrngghaprvkelllstlspeqlvltlleaeppnvlvsrp 240
            ||:||||||:  ||:||  :||||:||||||||:||||  :||||||||||||||||||:||:|||
Qy      226 YRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQLVLTLLEAEPPHVLISRP 285

Db      241 smpfteasmmmsltkladkelvhmigwakkipgfvelslldqvrllescwmevlmvglmw 300
            | ||||||||||||||||||||||||:||||||||||||||:|||||||||||||||||:|:||||
Qy      286 SAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMW 345

Db      301 rsidhpgklifapdlvldrdegkcvegileifdmllattsrfrelklqhkeylcvkamil 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      346 RSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMIL 405

Db      361 lnssmyplasanqeaessrklthllnavtdalvwviaksgissqqqsvrlanllmllshv 420
            ||||||||  :|  |:|:|||||:||||||||||||||||||||||||||:||||||||||||||
Qy      406 LNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHV 465

Db      421 rhisnkgmehllsmkcknvvpvydlllemlnahtlrgykssisgsecsstedsknkessq 480
            || |||||||||:||||||||||||||||||||||| |||  ||||:||||||::||||:||:||
Qy      466 RHASNKGMEHLLNMKCKNVVPVYDLLLEMLNAHVLRGCKSSITGSECSPAEDSKSKEGSQ 525

Db      481 nlqsq 485
            | |||
Qy      526 NPQSQ 530

RESULT    3
ID   W14725 standard; Protein; 485 AA.
AC   W14725;
DT   08-JUN-1997  (first entry)
DE   Mouse oestrogen receptor beta (ER-beta).
KW   Orphan receptor; oestrogen receptor beta; ER-beta;
KW   nuclear receptor; prostate cancer; benign prostatic hyperplasia;
KW   ovary cancer; cardiovascular disease; osteoporosis;
KW   environment; pollutant.
OS   Mus sp.
FH   Key             Location/Qualifiers
FT   Domain          104..169
FT   /label= DNA-binding_domain
FT   Domain          260..457
FT   /label= Ligand-binding_domain
PN   WO9709348-A2.
PD   13-MAR-1997.
PF   09-SEP-1996; E03933.
PR   08-SEP-1995; GB-018272.
PR   15-MAR-1996; GB-005550.
PR   11-APR-1996; GB-007532.
PR   08-MAY-1996; GB-009576.
PA   (KARO-) KARO BIO AB.
PI   Enmark E,  Gustafsson J,  Kuiper GG;
```

```
DR   WPI; 97-192842/17.
DR   N-PSDB; T62844.
PT   New isolated oestrogen receptor beta - used to develop prods. for
PT   treating e.g. cancers, CNS diseases, osteoporosis or cardiovascular
PT   disease
PS   Claim 1; Fig 14A; 45pp; English.
CC   A novel mouse orphan receptor (W14725) is related to the known
CC   oestrogen receptor ER-alpha, and has been designated ER-beta.
CC   It is an oestrogen receptor-related nuclear receptor. The ER-beta
CC   amino acid sequence was deduced from a cDNA clone (T62844).
CC   Rat, human and mouse ER-beta (W14723-25) can be used to isolate
CC   molecules for use in the treatment of cardiovascular diseases,
CC   central nervous system diseases, osteoporosis, prostate or ovarian
CC   cancer or benign prostatic hyperplasia and to test environmental
CC   chemicals for oestrogenic activity.
SQ   Sequence   485 AA;

  Query Match              93.2%;  Score 2250;  DB 21;  Length 485;
  Best Local Similarity    88.0%;  Pred. No. 0.00e+00;
  Matches  427;  Conservative   36;  Mismatches   22;  Indels    0;  Gaps    0;

Db        1 mafyspavmnysvpsstgnleggpvrqtaspnvlwptsghlsplathcqssllyaepqks 60
            |:|||||||||||.||.   |||||| |||:||||||||:||||||    | | | ||||||||
Qy       46 MTFYSPAVMNYSIPSNVTNLEGGPGRQTTSPNVLWPTPGHLSPLVVHRQLSHLYAEPQKS 105

Db       61 pwcearslehtlpvnretlkrklggsgcaspvtspstkrdahfcavcsdyasgyhygvws 120
            |||||||||||||:|||||||||::|: ||||||:|::|||||||||||||||||||||||
Qy      106 PWCEARSLEHTLPVNRETLKRKVSGNRCASPVTGPGSKRDAHFCAVCSDYASGYHYGVWS 165

Db      121 cegckaffkrsiqghndyicpatnqctidknrrkscqacrlrkcyevgmvkcgsrrercg 180
            |.||||||||||||||||||||||'||||||||:||||||||||||||||||||||||||||
Qy      166 CEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCG 225

Db      181 yrivrrqrsaseqvhclnkakrtsghtprvkelllnslspeqlvltlleaeppnvlvsrp 240
            ||:||'||||| ||:||'  ||||::||:|||:||||::'||||||||||||||||:||:|||
Qy      226 YRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQLVLTLLEAEPPHVLISRP 285

Db      241 smpfteasmmmsltkladkelvhmigwakkipgfvelslldqvrllescwmevlmvglmw 300
            | ||||||||||||||||||||||||||:||||||||||||||:||||||||||||||:||||
Qy      286 SAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMW 345

Db      301 rsidhpgklifapdlvldrdegkcvegileifdmllattarfrelklqhkeylcvkamil 360
            |||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Qy      346 RSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMIL 405

Db      361 lnssmyhlatasqeaessrklthllnavtdalvwvisksrissqqqsvrlanllmllshv 420
            ||||||  |  ||:|:|:||||||:|||||||||||||||||:||  ||||||||:|||'|||||
Qy      406 LNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHV 465

Db      421 rhisnkgmehllsmkcknvvpvydlllemlnahtlrgykssisgsgccstedskskegsq 480
            || ||||||||'|:||||||||||||||:||||||||| ||| |||||·||  |  ··||||||||||
Qy      466 RHASNKGMEHLLNMKCKNVVPVYDLLLEMLNAHVLRGCKSSITGSECSPAEDSKSKEGSQ 525

Db      481 nlqsq 485
            | |||
Qy      526 NPQSQ 530


RESULT    4
ID   P70543 standard; protein; 595 AA
AC   P70543;
DT   24-FEB-1991  (first entry)
DE   Human oestrogen receptor protein.
KW   Human oestrogen receptor protein; expression system; assay;
OS   Homo sapiens.
PN   WO8705049-A.
PD   27-AUG-1987.
PF   18-FEB-1987; U00341.
PR   20-FEB-1986; US-833829.
PA   (CALB-) California Biotechnology Inc.
PI   Shine J;
DR   WPI; 87-250213/35
DR   N-PSDB; N70880
PT   Expression system for vertebrate steroid receptor protein - comprising
PT   DNA sequence encoding the protein linked to control sequences in
PT   eucaryotic hosts.
PS   Disclosure; Fig. 1-1 - 1-2; 26pp; English.
CC   steroid receptor protein;
CC   The protein is expressed in an expression system from cDNA operably
CC   linked to control sequences compatible with eukaryotic host cells.
CC   This method allows expression under conditions which favour appropriate
CC   post-translational processing.  It also produces large amts. of purified
CC   protein useful in the design of agonist and antagonist cpds. for study
CC   of the mechanism of action of the steroid binding proteins in general,
CC   and for use in diagnostic assays for the proteins or antibodies to
CC   them.  These assays are important in, eg the diagnosis of tumour
CC   sensitivities to steroid metabolism.  Suitable host cells are VERO,
CC   HeLa and CHO cells.
SQ   Sequence   595 AA;

  Query Match              42.9%;  Score 1671;  DB 2;  Length 595;
  Best Local Similarity    56.3%;  Pred. No. 3.53e-151;
  Matches  227;  Conservative   98;  Mismatches   61;  Indels   17;  Gaps   13;

Db      179 saketrycavcndyasgyhygvwscegckaffkrsiqghndymcpatnqctidknrrksc 237
            | :::::||||:||||||||||||||||||||||||||||||||:||||||'||||||||||'
Qy      142 SKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSC 201

Db      238 qacrlrkcyevgmmkggirkdrrggrmlkhkrqrddgegrgevgsagdmraanlwpsplm 297
            |||||||||||||:| | |::| | |:::  |||    | :    :: : :|::     :|
Qy      202 QACRLRKCYEVGMVKCGSRRERCGYRLVR--RQRSADE-QLHCAGKA-KRSGG--HAP-R 254

Db      298 ikrskknslalsltadqmvsalldaepp-ilyseydptrpfseasmmglltnladrelvh 356
            : |       |   :|::::|:' :||:||||  :| |    |: ||:|||||  ||:|||:||||'
Qy      255 V-RELL--LD-ALSPEQLVLTLLEAEPPHVLISR--PSAPFTEASMMMSLTKLADKELVH 308

Db      357 minwakrvpgfvdltlhdqvhllecawleilmiglvwrsmehpgkllfapnllldrnqgk 416
            ||:|||::||||:|:| ||||:|||  |:|:||:||:|||::|| ||:|||':|:||||::||
Qy      309 MISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGK 368

Db      417 cvegmveifdmllatssrfrmmnlqgeefvclksiillnsgvytflsstlksleekdhih 476
            ||||  |||||||||:||||    ||    |  | .|.|  .||||||   |  · ··|   ·   ·     ·
Qy      369 CVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSRKLA 427

Db      477 rvldkitdtlihlmakagltlqqqhqrlaqlllilshirhmsnkgmehlysmkcknvvpl 536
            ::|: :||:|: ::||: :: |||  ||| ||::||':|| ||||||||| :||'||:·:
Qy      428 HLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPV 487

Db      537 ydlllemldahrlha-ptsrggasveetdqshlatagstsshs 578
            ||||||||:|| |::  :|  |:     :::|  :  || :::|
Qy      488 YDLLLEMLNAHVLRGCKSSITGSECSPAEDSK-SKEGSQNPQS 529


RESULT    5
ID   P80925 standard; protein; 334 AA.
AC   P80925;
DT   29-DEC-1990  (first entry)
DE   Sequence of the human oestrogen receptor
KW   Hormone receptor; hormone-binding; transcription activation;.
OS   Homo sapiens.
PN   WO8803168-A.
PD   05-MAY-1988.
PF   23-OCT-1987; U02782.
PR   20-OCT-1987; US-108471.
PA   (SALK) Salk Inst for Biol Stud.
PI   Evans RM, Weinberger CA, Hollenberg SM, Giguere V;
DR   WPI; 88-133242/19.
PT   Recombinant DNA encoding hormone recpetors -
PT   comprising glucocorticoid, mineralocorticoid, thyroid hormone
PT   and novel hormone receptors
PS   Example; Fig III-2; 243pp; English.
CC   DNAs encoding hormone receptors and the hormone receptors themselves are
CC   claimed. The DNA can be used to make the hormone receptor proteins and
CC   functional modified forms in quantities not previously possible. The
CC   receptor proteins can be used to screen cpds. for receptor-agonist or
CC   receptor-antagonist activity. They can also be used in diagnostic assays.
```

```
SQ   Sequence   334 AA;

  Query Match             36.5%;  Score 1424;  DB 2;  Length 334;
  Best Local Similarity   57.4%;  Pred. No. 1.87e-126;
  Matches  193;  Conservative   80;  Mismatches   47;  Indels   16;  Gaps   12;

Db          1 ycavcndyasgyhygvwsceqckaffkrs-gqhndymcpatnqctidknrrksaqacrl: 59
              :||||:||||||||||||||||||||||| ||||||:||||||||||||||||||||||
Qy        148 FCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCIIDKNRRKSCQACRLR 207

Db         60 kcyrvgmmkqgirkdrrgqrmlkhkrqrddgeqrgevqsaqdmraanlwpsplmikrskk 119
              ||| |||:| | |::| | |:::  |||    | :    :: : :|::     :|  : |
Qy        208 KCYEVGMVKCGSRRERCGYRLVR  RQRSADE-QLHCAGKA-KRSGG--HAP-RV-RELL 259

Db        120 nslalsltadqmvsalldaepp-ilyseydptrpfseasmmglltnladrelvhminwak 178
                 |   :|::::|:| :||:|||| :| |    |: ||:|||||   ||:|||:||||||:|||
Qy        260 --LD-ALSPEQLVLTLLEAEPPHVLISR--PSAPFTEASMMMSLTKLADKELVHMISWAK 314

Db        179 rvpgfvdltlhdqvhllecawlcilmiglvwrsmehpvkllfapnllldrnqgkcvegmv 238
              ::|||||:|:| |||:|||  |:|:||:||:|||::|| ||:|||:|:|||::||||||::
Qy        315 KIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGIL 374

Db        239 eifdmllatsssrfrmmnlqgeefvclksiillnsgvytflsstlksleekdhihrvldk 298
              ||||||||:||||  :.||  |..|:|..|||||..| ....|   .   .     . ..|. .
Qy        375 EIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSRKLAHLLNAV 433

Db        299 tdtlihlmakagltlqqqhqrlaqllliilshirhms 334
              ||:|: ::||:|:: |||  ||| ||::|||:|| |
Qy        434 TDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHAS 469


RESULT     6
ID   P80931 standard; protein; 433 AA.
AC   P80931;
DT   29-DEC-1990  (first entry)
DE   Sequence of human estrogen related receptor protein (hERR2)
KW   Hormone receptor, hormone binding, transcription activation.
OS   Homo sapiens.
PN   WO8803168-A.
PD   05-MAY-1988.
PF   23-OCT-1987; U02782.
PR   20-OCT-1987; US-108471.
PA   (SALK) Salk Inst for Biol Stud.
PI   Evans RM, Weinberger CA, Hollenberg SM, Giguere V.
DR   WPI; 88-133242/19.
DR   N-PSDB; N80922.
PT   Recombinant DNA encoding hormone receptors -
PT   comprising glucocorticoid, mineralocorticoid, thyroid hormone
PT   and novel hormone receptors
PS   Claim 17; Fig V-2(B)-1 and -2; 243pp; English.
CC   DNAs encoding hormone receptors and the hormone receptors themselves are
CC   claimed. The DNA can be used to make the hormone receptor proteins and
CC   functional modified forms in quantities not previously possible. The
CC   receptor proteins can be used to screen cpds. for receptor-agonist or
CC   receptor-antagonist activity. They can also be used in diagnostic assays.
CC   Also claimed is pure DNA encoding oestrogen-related receptors hERR1 and
CC   hERR2 and hERR1 and hERR2. The new hERR1 and hERR2 receptors will provide
CC   the basis for development of an assay system that will lead to the
CC   identification of novel hormones.
SQ   Sequence   433 AA;

  Query Match             20.7%;  Score 807;  DB 2;  Length 433;
  Best Local Similarity   39.5%;  Pred. No. 3.05e-65;
  Matches  140;  Conservative   84;  Mismatches  108;  Indels   22;  Gaps   14;

Db        101 rlclvcgdiasgyhygvasceackaffkrtiqgnieysepatneceitkrrrksgqacrf 160
              ::| ||:| |||||||| |||:|||||||||:|||: :| |||||:| | | ||||||||:
Qy        147 HFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCIIDKNRRKSCQACRL 206

Db        161 mkclkvgmlkegvrldrv  r gqrqky krrldsenspylslqisppakkpltkivs-- 214
               ||  |||:| | | :|    |   ||:     :|       :  |   :|  :  |   :|
Qy        207 RKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPE 266
```

```
Db        215 yl-l-vacpdklyam-ppddvpeqdlkalttlcdladrelvflliswakhipgfsnltlgd 271
               | | : |::    : :    |  :    : :|  |||:||| :||||| |||| :|:| |
Qy        267 QLVLILLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 326

Db        272 qmsllqsawmeililgivyrslpyddklayaedyimdeehsrlv qllelyrailqlvrr 330
              |: ||:| |||:!::|:: ||:     || :| | ::| : :: | |:|||::  :|     |
Qy        327 QVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSR 386

Db        331 ykklkvekeefvmlkalalansdsmy--i---enleav qklqdllh-  ealq dyels- 380
              :: ||::  |:: :||: | || |||   :     :: :: :||  ||:    :||      |
Qy        387 FRELKLQHKEYLCVKAMILLNS-SMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKSG 445

Db        381 -qrheeqrragklltlpllrqtaakavqhfysvklqgkvpmhklflemleakv 433
                 :::: | ::||: |: :|::: |::|: ::|     ||:  |:||||:| |
Qy        446 ISSIQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNAHV 499


RESULT     7
ID   P80930 standard; protein; 521 AA.
AC   P80930;
DT   10-MAR-1993  (revised)
DT   29-DEC-1990  (first entry)
DE   Sequence of human estrogen related receptor protein (hERR1)
KW   Hormone receptor, hormone-binding, transcription activation.
OS   Homo sapiens.
PN   WO8803168-A.
PD   05-MAY-1988.
PF   23-OCT-1987; U02782.
PR   20-OCT-1987; US-108471.
PA   (SALK) Salk Inst for Biol Stud.
PI   Evans RM, Weinberger CA, Hollenberg SM, Giguere V;
DR   WPI; 88-133242/19.
DR   N-PSDB; N80921.
PT   Recombinant DNA encoding hormone receptors -
PT   comprising glucocorticoid, mineralocorticoid, thyroid hormone
PT   and novel hormone receptors
PS   Claim 17; Fig V-1(B)-1 and -2; 243pp; English.
CC   DNAs encoding hormone receptors and the hormone receptors themselves are
CC   claimed. The DNA can be used to make the hormone receptor proteins and
CC   functional modified forms in quantities not previously possible. The
CC   receptor proteins can be used to screen cpds. for receptor agonist or
CC   receptor-antagonist activity. They can also be used in diagnostic assays.
CC   Also claimed is pure DNA encoding oestrogen related receptors hERR1 and
CC   hERR2 and hERR1 and hERR2. The new hERR1 and hERR2 receptors will provide
CC   the basis for development of an assay system that will lead to the
CC   identification of novel hormones.
SQ   Sequence   521 AA;

  Query Match             20.5%;  Score 800;  DB 2;  Length 521;
  Best Local Similarity   39.9%;  Pred. No. 1.49e-64;
  Matches  141;  Conservative   74;  Mismatches  128;  Indels   10;  Gaps    7;

Db        174 rlclvcgdvasgyhygvasceackaffkrtiqgsieysepasneceitkrrrkacqacrf 233
              ::| ||:| |||||||| |||:|||||||||:|||  :| |||:|:| | | |||::||||:
Qy        147 HFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRL 206

Db        234 tkclrvgmlkeqvrldrvrggrqkykrrpevdplpfpg p-fpagplavagprktaapv 291
               ||  |||.| | | .|   | |    ::|:   : |   :| :   ::|   :        | :
Qy        207 RKCYEVGMVKCGSRRERC-GYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSP 265

Db        292 nalvshllvvepeklyampdpagpdghlpavatlcdlfdrelvvtiswaksipgfsslsl 351
               : ||  ||  ||  :  :: |::|      : : :|  | |:|:|  |||||  ||||   |||
Qy        266 EQLVLILLEAEPPHVL-ISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSL 324

Db        352 sdqmsvlqsvwmevlvlgvaqrslplqdelafaedlvld-eegaraaglqelqaallqlv 410
               |. .|.| ||||::|:  ||:       | || |||| :||      |: |:    ||
Qy        325 FDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATT 384

Db        411 rrlqalrlereeyvllkalalansdsvhiedeprlwsscekllh--eal---leyeagra 465
               |·· |·|:: ||: :||: | ||     :      :    |  || |   :|:    :    :::
Qy        385 SRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKS 444

Db        466 gpqggaerrraqrllltlpllrqtagkvlahfyqvklegkvpmhklflemlea 518
```

```
             |::  :  | : ||: |: :|::: | : |:  :|     ||:  ::'|||:|
Qy    445 GISGQQQSMRLANLIMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNA 497


RESULT    8
ID    P80919 standard; protein; 777 AA.
AC    P80919;
DT    29-DEC-1990  (first entry)
DE    Sequence of the primary protein sequence of human glucocorticoid
DE    receptor (hGR)
KW    Hormone receptor; hormone-binding; transcription activation.
OS    Homo sapiens
PN    WO8803168-A.
PD    05-MAY-1988.
PF    23-OCT-1987; U02782.
PR    20-OCT-1987; US-108471.
PA    (SALK) Salk Inst for Biol Stud.
PI    Evans RM, Weinberger CA  Hollenberg SM  Giguere V;
DR    WPI; 88-133242/19.
DR    N-PSDB; N80916
PT    Recombinant DNA encoding hormone recpetors -
PT    comprising glucocorticoid, mineralocorticoid, thyroid hormone
PT    and novel hormone receptors
PS    Claim 24; Fig I-2(1) to (2); 243pp  English
CC    DNAs encoding hormone receptors and the hormone receptors themselves are
CC    claimed. The DNA can be used to make the hormone receptor proteins and
CC    functional modified forms in quantities not previously possible. The
CC    receptor proteins can be used to screen cpds  for receptor-agonist or
CC    receptor-antagonist activity  They can also be used in diagnostic assays.
SQ    Sequence   777 AA;

  Query Match            15.7%;  Score 614;  DB 2;  Length 777;
  Best Local Similarity  33.4%;  Pred. No. 2.11e-46;
  Matches   97;  Conservative   84;  Mismatches  93;  Indels  16;  Gaps  15;

Db    397 smrpdvssppsssstattgp-p-pklclvcsdeasgchygvltcgsckvffkravegqhn 454
            :::  ||:   :|: :  |:    : :| |||| ||| |||| :| :|| ||||:::|:::
Qy    123 ILPFFVSGNPIASPVIGHGSPPLAHFIAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHND 182

Db    455 ylcagrndciidkirrkncpacryrkclqagmnl-earktk-k-kikgiqqattg-v-sq 509
          |:|:: |:| ||| |||:| ||| ||| : ||       :|: :    ::    |:::    :
Qy    183 YICPATNQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCA 242

Db    510 etsenpgnkt-ivpatlpq-ltpt-lv-sllevieprvlyagydssvpdstwrimttlnm 565
              :  :|  :  |   | : |:|  || :|||   | ||  .  ..  ...  .| .| :
Qy    243 GKAKRSGGHAPRVRELLLDALSPEQLVLTLLEAEPPHVLISRPSAPFTEAS--MMMSLTK 300

Db    566 lggrqviaavkwakaipgfrnlhlddqmtllqyswmflmafalgwrsyrqssanllcfap 625
          |: ::::  : ||| |||| :| | ||: ||: ||       | ||| ·     |· |||
Qy    301 LADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPG-KLI-FAP 358

Db    626 dliin-eqrmtlpcmydqckhmlyvsselhrlqvsyeeylcmktllllss 674
          |':::: :: :  :   :  :     :|  :| :: | :   ||||:|:::||:|
Qy    359 DLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNS 408


RESULT    9
ID    R12224 standard  Protein  902 AA
AC    R12224;
DT    20-AUG-1991  (first entry)
DE    Rat androgen receptor.
KW    rAR; DNA-binding protein; steroid hormone.
OS    Rattus rattus
FH    Key                  Location/Qualifiers
FT    Domain               540..610
FT    /label= DNA-binding domain
FT    /note= "cysteine-rich"
PN    WO9107423-A
PD    30-MAY-1991.
PF    19-OCT-1990; U06015.
PR    17-NOV-1989; US-438775.
PA    (ARCH-) ARCH DEV CORP.
PI    Liao S. Chang C;
DR    WPI; 91-178048/24.
DR    N-PSDB; [illegible]
PT    Androgen receptor and TR2 DNA binding proteins - DNA sequences
PT    and antibodies for detection and quantification methods
PS    Claim 25; Fig 3; 79pp; English.
CC    This sequence was deduced from a cDNA clone isolated by screening
CC    a rat ventral prostate lambda gt11 library in E.coli Y1090. The
CC    sequence is very similar to that of human AR and in
CC    the DNA-binding domain it is identical to that of hAR DNA-binding
CC    domain  Homology comparisons with other known steroid receptors
CC    indicate that rAR is more closely related to glucocorticoid,
CC    mineralo-corticoid and progesterone receptors than to v-erb A or to
CC    receptors for oestrogen, vitamin D and thyroid hormones.
SQ    Sequence   902 AA;

  Query Match            14.8%;  Score 576;  DB 3;  Length 902;
  Best Local Similarity  30.1%;  Pred. No. 1.00e-42;
  Matches  108;  Conservative  103;  Mismatches 126;  Indels  22;  Gaps  19;

Db    542 clicgdeasgchygaltcgsckvffkraaegkqkylcasrndctidkfrrkncpscrlrk 601
           | :|:| ||| |||   :| :|| ||||: :|    |:|:: |:||||| |||:| :||||
Qy    149 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRLRK 208

Db    602 cyeagmtl-garklk--klgnlklqeegenssagsptedpsqkmtvshi--egyecqpi 655
          ||| ||   |:|:  .  .:        . :   ||  .    ::    |  :  ::   : :
Qy    209 CYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 268

Db    656 flnvleaiepgvvcaghdnnqpdsfaallsslnelgerqlvhvvkwakalpgfrnlhvdd 715
            | .|||  | |.  .. .  | : |... ||  |....|||.. ||| .'|| :| : |
Qy    269 VLTLLEAEPPHVL-ISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 326

Db    716 qmaviqyswmglmvfamgwrsftnvnsrmlyfapdlvfn-eyrmhksrmysqcvrmrhls 774
          |. ..| ||  ... .. |||. .  :::: |||||||:: :  :    :      :  :
Qy    327 QVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDMLLATT 384

Db    775 qefgwlqitpqeflcmkal-ll-fsiip-vdglknqkffdel-rm-nyik-eldriiack 828
            |  | :   |:||:'|: ||  |: | | :  :       | :: | : : | :::|
Qy    385 SRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKS 444

Db    829 rknptscsrrfyqltklldsv-qpiarelhqftfdllikshmvsvdfpemmaeiisvqv 886
             ::   | |:  | .|| | ..  . .... ...  | ..|:| : ::: |::: :|
Qy    445 GISGQQQSMRLANLLMLLSHVRHASNKGMEHL-LNMKCK-NVVPV-Y-DLLLEMLNAHV 499


RESULT   10
ID    P93110 standard; protein; 902 AA.
AC    P93110;
DT    19-MAR-1990  (first entry)
DE    Rat androgen receptor.
KW    Rat androgen receptor; monoclonal antibody; polyclonal antibody; cancer.
OS    Rattus rattus.
PN    WO8909791-A.
PD    19-OCT-1989.
PF    13-APR-1989; U01548.
PR    14-APR-1988; US-182646.
PA    (UYNC-) University of North Carolina.
PI    French FS, Wilson EM, Joseph DR, Lubahn DB;
DR    WPI; 89-324206/44.
DR    N-PSDB; N91773.
PT    DNA encoding androgen receptor protein - useful for transforming
PT    eukaryotic hosts for protein expression and subsequent antibody prodn.
PS    Disclosure; Fig. 5; 41pp; English.
CC    Androgen receptor protein (AR) is used to produce mono- or poly clonal
CC    antibodies.  These are used for the detection and quantification of AR in
CC    the presence of endogenous androgen, as androgen will not interfere with
CC    binding.  They may be used in assays to determine and quantify cellular
CC    distribution of AR in tumour tissue, and are esp. useful for evaluating
CC    prostate cancers to determine responsiveness to androgen withdrawal
CC    therapy.
SQ    Sequence   902 AA;

  Query Match            14.6%;  Score 570;  DB 1;  Length 902;
  Best Local Similarity  29.8%;  Pred. No. 3.81e-42;
  Matches   107;  Conservative  103;  Mismatches 127;  Indels  22;  Gaps  19;

Db     542 cliqdeasgchygaltcgsckvffkraaegkqkylcasrndctidkfrrkncpscrlrk 601
           | :|:| ||| |||   :| :|| ||||: :|    |:|:: |:||||| |||:| :|||||
Qy     149 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 208

Db     602 cyeagmtl-garklk---klgnlklqeegeassagsptedpsqkmtvshi--egyecqpi 655
           ||| ||    |:|: :    :|          : :    || :      ::    |  :  ::   : :
Qy     209 CYEVGMVKCGSRRERCGYRLVRRQRSAEEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 268

Db     656 flnvleaiepgvvcaghdnnqpdsfaallsslnelgerqlvhvvkwakalpgfrnlhvdd 715
             | :|||  | |:   :: :   | : |::: ||   |::::|||:: ||| :||| :| : |
Qy     269 VLTLLEAEPPHVL-ISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 326

Db     716 qmaviqyswmglmvfamgwrsftnvnsrmlyfapdlvfn-eyrmhksrmysqcvrmrhls 774
           |: :::   || ::: :: |||: :   :::: |||||||:: :  :      :       :  :
Qy     327 QVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDMLLATT 384

Db     775 qefgwlqitpqeflcmkal-ll-fsiip-vdglknqkffdel-rm-nyike-ldriiack 828
              |  | :   |:||.||. ||  |. | | .   .         | .. | . . | ..||
Qy     385 SRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKS 444

Db     829 rknptscsrrfyqltklldsv-qpiarelhqftfdllikshmvsvdfpemmaeiisvqv 886
                ..    | |.  | .||  | ..  . . . . .     |    | |         :      |
Qy     445 GISSQQQSMRLANLLMLLSHVRHASNKGMEHL-LNMKCK-NVVPV-Y-DLLLEMLNAHV 499


RESULT   11
ID   P91006 standard; protein; 902 AA.
AC   P91006;
DT   28-FEB-1990  (first entry)
DE   Rat androgen receptor DNA clone.
KW   Androgen receptor; TR2 polypeptide.
OS   Rat
FH   Key             Location/Qualifiers
FT   Region          1..902
FT   /*tag=  a
FT   /product=98 KD polypeptide
FT   Region          170..902
FT   /*tag=  b
FT   /product=79 KD polypeptide
PN   W08909223-A.
PD   05-OCT-1989.
PF   24-MAR-1989; U01238.
PR   30-MAR-1988; US-176107.
PA   (ARCH-) Arch Development Corp.
PI   Liao S, Chang C;
DR   WPI; 89-309501/42.
DR   N-PSDB; N91578.
PT   New DNA encoding new androgen receptor and TR2 polypeptide(s) - able
PT   to bind DNA, and derived antibodies, useful for receptor assay and
PT   purification.
PS   Claim 8; fig. 3; 60pp; English.
CC   The protein is used to raise antibodies for receptor assays and for
CC   affinity purification.
CC   The 98 kD product starts at the first Met codon, the 79 kD product
CC   starts from the second.
SQ   Sequence   902 AA;

  Query Match              14.6%;  Score 570;  DB 1;  Length 902;
  Best Local Similarity  29.8%;  Pred. No. 3.81e-42;
  Matches   107;  Conservative  103;  Mismatches 127;  Indels  22;  Gaps  19;

Db     542 cliqdeasgchygaltcgsckvffkraaegkqkylcasrndctidkfrrkncpscrlrk 601
           | :|:| ||| |||   :| :|| ||||: :|    |:|:: |:||||| |||:| :|||||
Qy     149 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 208

Db     602 cyeagmtl-garklk---klgnlklqeegeassagsptedpsqkmtvshi--egyecqpi 655
           ||| ||    |:|: :    :|          : :    || :      ::    |  :  ::   : :
Qy     209 CYEVGMVKCGSRRERCGYRLVRRQRSAEEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 268

Db     656 flnvleaiepgvvcaghdnnqpdsfaallsslnelgerqlvhvvkwakalpgfrnlhvdd 715
             | :|||  | |:   :: :   | : |::: ||   |::::|||:: ||| :||| :| : |
Qy     269 VLTLLEAEPPHVL-ISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 326

Db     716 qmaviqyswmglmvfamgwrsftnvnsrmlyfapdlvfn-eyrmhksrmysqcvrmrhls 774
           |: :::   || ::: :: |||: :   :::: |||||||:: :  :      :       :  :
Qy     327 QVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDMLLATT 384

Db     775 qefgwlqitpqeflcmkal-ll-fsiip-vdglknqkffdel-rm-nyike-ldriiack 828
              |  | :   |:||:|||: ||  |: | | :   :         | :: | : : | ::||
Qy     385 SRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKS 444

Db     829 rknptscsrrfyqltklldsv-qpiarelhqftfdllikshmvsvdfpemmaeiisvqv 886
                ::    | |:  | :||  | ::  : : :: :::   | ::|:| : ::: |::: :|
Qy     445 GISSQQQSMRLANLLMLLSHVRHASNKGMEHL-LNMKCK-NVVPV-Y-DLLLEMLNAHV 499


RESULT   12
ID   R12223 standard; Protein; 918 AA.
AC   R12223;
DT   20-AUG-1991  (first entry)
DE   Human androgen receptor.
KW   hAR; DNA-binding protein; steroid hormone.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   Domain          556..626
FT   /label= DNA-binding domain
FT   /note= "cysteine-rich"
PN   WO9107423-A.
PD   30-MAY-1991.
PF   19-OCT-1990; U06015.
PR   17-NOV-1989; US-438775.
PA   (ARCH-) ARCH DEV CORP.
PI   Liao S, Chang C;
DR   WPI; 91-178048/24.
DR   N-PSDB; Q12001.
PT   Androgen receptor and TR2 DNA binding proteins - DNA sequences
PT   and antibodies for detection and quantification methods
PS   Claim 25; Fig 3; 79pp; English.
CC   This sequence was deduced from a cDNA clone isolated by screening
CC   commercially available human testis and prostate lambda gt11 cDNA
CC   libraries. The sequence is very similar to that of rat AR and in
CC   the DNA-binding domain it is identical to that of rAR DNA-binding
CC   domain. Homology comparisons with other known steroid receptors
CC   indicate that hAR is more closely related to glucocorticoid,
CC   mineralocorticoid and progesterone receptors than to v-erb-A or to
CC   receptors for oestrogen, vitamin D and thyroid hormones.
SQ   Sequence   918 AA;

  Query Match              14.6%;  Score 570;  DB 3;  Length 918;
  Best Local Similarity  29.5%;  Pred. No. 3.81e-42;
  Matches   107;  Conservative  106;  Mismatches 125;  Indels  22;  Gaps  20;

Db     558 cliqdeasgchygaltcgsckvffkraaegkqkylcasrndctidkfrrkncpscrlrk 617
           | :|:| ||| |||   :| :|| ||||: :|    |:|:: |:||||| |||:: :|||||
Qy     149 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 208

Db     618 cyeagmtl-garklk---klgelklqeeg-eassttsptee-tqkltvshi-eqy-qpi 671
           ||| ||    |:|: :    :|      :  :|       :: :   :     : ::      :::    : :
Qy     209 CYEVGMVKCGSRRERCGYRLVRRQRSAEEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 268

Db     672 flnvleaiepgvvcaghdnnqpdsfaallsslnelgerqlvhvvkwakalpgfrnlhvdd 731
             |  |||  | |:              |     |       ||   |::::|||:: ||| :||| :| : |
Qy     269 VLTLLEAEPPHVL-ISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 326

Db     732 qmavieyswmglmvfamgwrsftnvnsrmlyfapdlvfn-eyrmhksrmysqcvrmrhls 790
           |: ::|   || ::: :: |||: :   :::: |||||||:: :  :      :       :  :
Qy     327 QVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDMLLATT 384

Db     791 qefgwlqitpqeflcmkal-ll-fsiip-vdglknqkffdel-rm-nyike-ldriiack 844
              |  | :   |:||:|||: ||  |: | | :   :         | :: | : : | ::||
Qy     385 SRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKS 444

Db     845 rknptscsrrfyqltklldsv-qpiarelhqftfdllikshmvsvdfpemmaeiisvqv 902
                ::    | |:  | :||  | ::  : : :: :::   | ::|:| : ::: |::: :|
Qy     445 GISSQQQSMRLANLLMLLSHVRHASNKGMEHL-LNMKCK-NVVPV-Y-DLLLEMLNAHV 499


RESULT   13
ID   P90996 standard; protein; 919 AA.
AC   P90996;
DT   28-FEB-1990  (first entry)
DE   Human androgen receptor DNA clone.
KW   Androgen receptor; TR2 polypeptide;
OS   Homo sapiens.
FH   Key                 Location/Qualifiers
FT   Region              1..919
FT   /*tag=  a
FT   /product=98 kD polypeptide
FT   Region              185..919
FT   /*tag=  b
FT   /product=79 kD polypeptide
PN   WO8909223-A.
PD   05-OCT-1989.
PF   24-MAR-1989; U01238.
PR   30-MAR-1988; US-176107.
PA   (ARCH-) Arch Development Corp.
PI   Liao S, Chang C;
DR   WPI; 89-309501/42.
DR   N-PSDB; N91577.
PT   New DNA encoding new androgen receptor and TR2 polypeptide(s) - able
PT   to bind DNA, and derived antibodies, useful for receptor assay and
PT   purification.
PS   Claim 8; fig. 3; 60pp; English.
CC   The protein is used to raise antibodies for receptor assays and for
CC   affinity purification.
CC   The 98 kD product starts at the first Met codon; the 79 kD product
CC   starts from the second
SQ   Sequence   919 AA;

  Query Match              14.5%;  Score 567;  DB 1;  Length 919;
  Best Local Similarity  29.2%;  Pred. No. 7.42e-42;
  Matches   105;  Conservative  107;  Mismatches 125;  Indels  22;  Gaps  20;

Db     559 clicgdeasgchygaltcgsckvffkraaegkqkylcasrndctidkfrrkncpscrlrk 618
           |.:|.| ||| |||  .| .|| ||||. .|   |:|.. |:||||| |||:| :|||||
Qy     149 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 208

Db     619 cyeagmtl-garklk---klgn-lklqeeg-eassttspteettqkltvshiegyecqpi 672
           ||| ||   |:|: :   :|    :  :|    :: :  :   : ::   ::: : :
Qy     209 CYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 268

Db     673 flnvleaiepgvvcaghdnnqpdsfaallsslnelgerqlvhvvkwakalpgfrnlhvdd 732
            | :|||  | |:    : : | : |::: ||  | :::|||:: ||| :||| :| : |
Qy     269 VLTLLEAEPPHVL-ISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 326

Db     733 qmaviqyswmglmvfamgwrsftnvnsrmlyfapdlvfn-eyrmhksrmysqcvrmrhls 791
           |: ...  || ... .. |||.    .... ||||||:: :    :     :   :   :
Qy     327 QVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDMLLATT 384

Db     792 qefgwlqitpqeflcmkal-ll-fsiip-vdglknqkffdel-rm-nyike-ldriiack 845
             |  | .   |:||:||. ||  |. | |  .     :  | .. | : : |  :||
Qy     385 SRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKS 444

Db     846 rknptscsrrfyqltklldsv-qpiarelhqftfdllikshmvsvdfpemmaeiisvqv 903
               :: | |   |  ||  |                 |  .|.| : ::: |::: :|
Qy     445 GISSQQQSMRLANLLMLLSHVRHASNKGMEHL-LNMKCK-NVVPV-Y-DLLLEMLNAHV 499


RESULT   14
ID   W14783 standard; Protein; 919 AA.
AC   W14783;
DT   22-JUN-1997  (first entry)
DE   Androgen receptor.
KW   Androgen receptor; acidic fibroblast growth factor; aFGF;
KW   antisense; benign prostatic hyperplasia; prostate cancer; therapy
OS   Homo sapiens.
PN   WO9711170-A1.
PD   27-MAR-1997.
PF   20-SEP-1996; U15081.
PR   20-SEP-1995; US-004018.
PA   (WORC-) WORCESTER FOUND BIOMEDICAL RES.
PI   Zamecnik PA,
DR   WPI; 97-202879/18.
DR   N-PSDB; T63407.
PT   Oligonucleotide(s) antisense to human androgen receptor and acidic
PT   FGF genes - used to inhibit gene expression, for the treatment of
PT   benign prostatic hyperplasia
PS   Disclosure; Page 22-28; 51pp; English.
CC   Human androgen receptor (W14783) binds testosterone and, acting
CC   at the transcriptional level, regulates the growth of normal
CC   prostatic cells.  Antisense oligonucleotides (see also T63200,
CC   T63404-05) based on an androgen receptor cDNA clone (see also
CC   T63407) can be used to prevent androgen receptor gene expression,
CC   thereby inhibiting the growth or survival of prostatic cells for
CC   the treatment of benign prostatic hyperplasia and prostate cancer.
SQ   Sequence   919 AA;

  Query Match              14.5%;  Score 567;  DB 21;  Length 919;
  Best Local Similarity  29.2%;  Pred. No. 7.42e-42;
  Matches   105;  Conservative  107;  Mismatches 125;  Indels  22;  Gaps  20;

Db     559 clicgdeasgchygaltcgsckvffkraaegkqkylcasrndctidkfrrkncpscrlrk 618
           |.:|.| ||| |||  .| .|| ||||. .|   |:|.. |:||||| |||:| :|||||
Qy     149 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 208

Db     619 cyeagmtl-garklk---klgn-lklqeeg-eassttspteettqkltvshiegyecqpi 672
           ||| ||   |:|: :   :|    :  :|    :: :  :   : ::   ::: : :
Qy     209 CYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 268

Db     673 flnvleaiepgvvcaghdnnqpdsfaallsslnelgerqlvhvvkwakalpgfrnlhvdd 732
            | :|||  | |:    : : | : |::: ||  | :::|||:: ||| :||| :| : |
Qy     269 VLTLLEAEPPHVL-ISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 326

Db     733 qmaviqyswmglmvfamgwrsftnvnsrmlyfapdlvfn-eyrmhksrmysqcvrmrhls 791
           |: ...  || ... .. |||.    .... ||||||:: :    :     :   :   :
Qy     327 QVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDMLLATT 384

Db     792 qefgwlqitpqeflcmkal-ll-fsiip-vdglknqkffdel-rm-nyike-ldriiack 845
             |  | .   |:||:||. ||  |. | |  .     :  | .. | : : |  :||
Qy     385 SRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKS 444

Db     846 rknptscsrrfyqltklldsv-qpiarelhqftfdllikshmvsvdfpemmaeiisvqv 903
               :: | |   |  ||  |                 |  .|.| : ::: |::: :|
Qy     445 GISSQQQSMRLANLLMLLSHVRHASNKGMEHL-LNMKCK-NVVPV-Y-DLLLEMLNAHV 499


RESULT   15
ID   P93109 standard; protein; 919 AA.
AC   P93109;
DT   19-MAR-1990  (first entry)
DE   Human androgen receptor.
KW   Human androgen receptor, ploycional antibody, cancer.
OS   Homo sapiens.
PN   WO8909791-A.
PD   19-OCT-1989.
PF   13-APR-1989; U01548.
PR   14-APR-1988; US-182646.
PA   (UYNC-) University of North Carolina.
PI   French FS, Wilson EM, Joseph DR, Lubahn DB,
DR   WPI; 89-324206/44.
DR   N-PSDB; N91772.
PT   DNA encoding androgen receptor protein - useful for transforming
PT   eukaryotic hosts for protein expression and subsequent antibody prodn.
PS   Disclosure; Fig. 4; 41pp; English.
CC   Androgen receptor protein (AR) is used to produce mono- or poly-clonal
CC   antibodies.  These are used for the detection and quantification of AR in
CC   the presence of endogenous androgen, as androgen will not interfere with
CC   binding.  They may be used in assays to determine and quantify cellular
CC   distribution of AR in tumour tissue, and are esp. useful for evaluating
CC   prostate cancers to determine responsiveness to androgen withdrawal
CC   therapy.
SQ   Sequence   919 AA;

  Query Match             14.5%;  Score 566;  DB 1;  Length 919;
  Best Local Similarity   29.2%;  Pred. No. 9.27e-42;
  Matches  105;  Conservative  107;  Mismatches 125;  Indels  22;  Gaps  20;

Db     559 elicqdeasqchygaltcqsckvffkraaegkqkylcasrndctidkfrrkncpscrlrk 618
           | :|:| ||| |||   :| :|| ||||: :|    |:|:: |:||||| |||:| :|||||
Qy     149 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRLRK 208

Db     619 cyeaqmtl-qarklk---klqn-lklqeeq-easttsptecttqkltvshiegyecqai 672
           ||| ||   |:|: :    :|     :  :|     :: :  :    : ::    :::   : :
Qy     209 CYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 268

Db     673 flnvleaiepavvcaghdnnqpdsfaallsslnelgerqlvhvvkwakalpgfrnlhvdd 732
             | :|||  | |:  :: :  | : |:.. ||  |....|||.. ||| :||| :| : |
Qy     269 VLTLLEAEPPHVL-ISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 326

Db     733 qmaviqyswmglmvfamgwrsftnvnsrmlyfapdlvfn-eyrmhksrmysqcvrmrhls 791
           |: :::  || ::: :: |||: :   :::: |||||||:: :  :     :       :   :
Qy     327 QVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDMLLATT 384

Db     792 qefqwlqitpqeflcmkal-ll-fsiip-vdglkngkffdel-rm-nyike-ldriiack 845
             |  | :   |:|||:||: ||  |: | | :  :       | :: | : : | ::||
Qy     385 SRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKS 444

Db     846 rkeptsesrrfyqltklldsv-qpiarelhqftfdllikshmvsvdfpemmaeiisvqv 903
             ::    | |:  | :||  | ::   : : :: :::  | ::|:| : ::: |::: :|
Qy     445 GISSQQQSMRLANLLMLLSHVRHASNKGMEHL-LNMKCK-NVVPV-Y-DLLLEMLNAHV 499

Search completed: Thu Dec 18 09:15:31 1997
Job time : 71 secs.
```

```
     ******************************************************************
                             MPSRCH (TM)
     ******************************************************************

              Release 2 1D John F. Collins, Biocomputing Research Unit
              Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
                     Distribution rights by IntelliGenetics, Inc.

MPsrch_nn   n.a. - n.a. database search, using Smith-Waterman algorithm

Run on:          Thu Dec 18 18:54:12 1997; MasPar time 140.57 Seconds
                                           928.836 Million cell updates/sec
Tabular output not generated.

Title:           >US-08-826-361A-2
Description:     (1-1251) from US08826361A.seq
Perfect Score:   1251
N.A. Sequence:      1 ATGAATTACAGCATTCCTAG         ACCTCAGGCATCCGAGTGA 1251
     Comp:            TACTTAATGTCGTAAGGATC         TGGAGTCCGTAGGCTCACT

Scoring table:   TABLE default
                 Gap 6

Nmatch   STD :   Dbase 0;  Query 0

Searched:        142090 seqs, 52183452 bases x 2

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database:        n-geneseq28
                   1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
                   8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
                   14:part14 15:part15 16:part16 17:part17 18:part18
                   19:part19 20:part20 21:part21 22:part22 23:part23
                   24:part24 25:part25 26:part26 27:part27 28:part28
                   29:part29

Statistics:      Mean 9.222   Variance 5.126,  scale 1.799

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution

                                  SUMMARIES
                  %
Result          Query
  No    Score   Match Length DB  ID            Description           Pred. No
------------------------------------------------------------------------------
    1    1247    99.7   1460 29  T62843        Human oestrogen recep  0.00e+00
    2     861    68.8   2568 29  T62842        Rat oestrogen recepto  0.00e+00
    3     843    67.4   1458 29  T62844        Mouse oestrogen recep  0.00e+00
    4     135    10.8   2092  2  N70980        cDNA encoding human b  2.37e-76
    5     135    10.8   4963 17  T06873        PhCMV*-1 promoter and  2.37e-76
    6     135    10.8   4963 13  Q76270        PhCMV*-1-controlled p  2.37e-76
    7     127    10.2    204 14  Q85976        Human estrogen recept  1.64e-70
    8      91     7.3   1047  2  Q10572        Human Natriuretic Pep  1.24e-44
c   9      71     5.7   3460 24  T42193        Fragment of pig oestr  1.13e-30
   10      68     5.4   2153  2  N90922        Sequence encoding hum  1.28e-28
c  11      66     5.3   1047  2  Q10572        Human Natriuretic Pep  2.95e-27
   12      63     5.0   6244 13  Q76269        PhCMV*-1-controlled p  3.17e-25
   13      63     5.0   6244 17  T06872        PhCMV*-1 promoter and  3.17e-25
   14      59     4.7   2992 13  Q79934        Human liver hap cDNA.  1.54e-22
   15      59     4.7   2992 11  T65572        Human liver hap cDNA.  1.54e-22
   16      59     4.7   2992 20  T06491        Human hepatoma retino  1.54e-22
   17      58     4.6   1912  2  Q10389        Encodes skin-specific  7.14e-22
   18      58     4.6   2521  3  Q20542        Mouse retinoic acid r  7.14e-22
   19      58     4.6   2741  3  Q20541        Mouse retinoic acid r  7.14e-22
   20      58     4.6   [illegible]  [illegible]        Sequence encoding the  7.14e-22
   21      56     4.5   1914 18  T05202        Murine retinoic acid   1.52e-20
   22      56     4.5   6176 27  T49226        Plasmid pGR0402B for   1.52e-20
   23      55     4.4   2989  1  N90093        Human Hap (hepatoma)   6.95e-20
   24      53     4.2   1231  6  Q37761        Sequence of a 1.23 kb  1.43e-18
   25      53     4.2   1795  2  Q12005        Human TR2-9 DNA bindi  1.43e-18
   26      53     4.2   2029  2  Q12004        Human TR2-5 androgen   1.43e-18
   27      53     4.2   2221  2  Q12006        Human TR2-11 DNA bind  1.43e-18
   28      53     4.2   2458  2  Q12003        Human TR2-7 DNA bindi  1.43e-18
   29      51     4.1   1576  2  Q10226        Clone pGEM-hFARgamma   2.90e-17
   30      51     4.1   1903  2  Q10399        Encodes human Retinoi  2.90e-17
   31      51     4.1   1903 18  T05196        Human retinoic acid r  2.90e-17
   32      51     4.1   3569  1  N91772        Human androgen recept  2.90e-17
   33      51     4.1   3569 29  T63407        Androgen receptor cDN  2.90e-17
   34      51     4.1   3715  2  Q12001        Full-length human and  2.90e-17
   35      50     4.0   1351 13  T72935        Human liver FAR beta   1.29e-16
   36      50     4.0   2249  2  N80921        Sequence encoding hum  1.29e-16
c  37      49     3.9    204  1  N81164        Base substituted E.co  5.72e-16
   38      47     3.8    387 19  T09863        Modified human glucoc  1.10e-14
   39      47     3.8    387 24  T45454        Modified human glucoc  1.10e-14
   40      47     3.8   3217  1  N91578        Rat androgen receptor  1.10e-14
   41      45     3.6   1893  2  Q12008        TrpE/androgen recepto  2.06e-13
   42      45     3.6   3017  2  Q12002        Full-length rat andro  2.06e-13
   43      45     3.6   3427 10  T56733        cDNA sequence encodin  2.06e-13
   44      45     3.6   4180  1  N91773        Rat androgen receptor  2.06e-13
   45      44     3.5   4400 18  T16151        Apoptopic cerebral ne  8.81e-13

                                  ALIGNMENTS

RESULT    1
ID   T62843 standard; cDNA; 1460 BP.
AC   T62843;
DT   08-JUN-1997  (first entry)
DE   Human oestrogen receptor beta (ER-beta) cDNA.
KW   Orphan receptor; oestrogen receptor beta; ER-beta;
KW   nuclear receptor; prostate cancer; benign prostatic hyperplasia;
KW   ovary cancer; cardiovascular disease; osteoporosis;
KW   environment; pollutant; ss.
OS   Homo sapiens.
FH   Key                Location/Qualifiers
FT   CDS                3..1460
FT   /*tag=  a
PN   WO9709348-A2.
PD   13-MAR-1997.
PF   09-SEP-1996, E03933.
PR   08-SEP-1995; GB-018272.
PR   15-MAR-1996; GB-005550.
PR   11-APR-1996; GB-007532.
PR   08-MAY-1996; GB-009576.
PA   (KARO-) KARO BIO AB.
PI   Enmark E,  Gustafsson J,  Kuiper GG;
DR   WPI; 97-191541/17.
DR   P-PSDB; W14724.
PT   New isolated oestrogen receptor beta - used to develop prods. for
PT   treating e.g. cancers, CNS diseases, osteoporosis or cardiovascular
PT   disease
PS   Claim 6; Fig 13B; 45pp; English.
CC   A cDNA clone (T62843) codes for a novel human oestrogen receptor-
CC   related nuclear receptor (W13724), designated ER-beta.  It was
CC   isolated from a human testis cDNA library using rat ER-beta cDNA
CC   (see also T62842) as probe.  The human ER-beta gene was mapped
CC   to chromosome 14 region 14q22-23.  Rat, human and mouse ER-beta
CC   (W14723-25) can be used to isolate molecules for use in the
CC   treatment of cardiovascular diseases, central nervous system
CC   diseases, osteoporosis, prostate or ovarian cancer or benign
CC   prostatic hyperplasia and to test environmental chemicals for
CC   oestrogenic activity.
SQ   Sequence  1460 BP;    352 A;    379 C;    411 G;    318 T;
```

```
  Query Match             99.7%;  Score 1247;  DB 29;  Length 1460;
  Best Local Similarity 100.0%;  Pred. No. 0.00e+00;
  Matches 1247;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Db       27 atgaattacagcattcccagcaatgtcactaacttggaaggtgggcctggtcggcagacc 86
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        1 ATGAATTACAGCATTCCCAGCAATGTCACTAACTTGGAAGGTGGGCCTGGTCGGCAGACC 60

Db       87 acaagcccaaatgtgttgtggccaacacctgggcacctttctcctttagtggtccatcgc 146
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       61 ACAAGCCCAAATGTGTTGTGGCCAACACCTGGGCACCTTTCTCCTTTAGTGGTCCATCGC 120

Db      147 cagttatcacatctgtatgcggaacctcaaaagagtccctggtgtgaagcaagatcgcta 206
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      121 CAGTTATCACATCTGTATGCGGAACCTCAAAAGAGTCCCTGGTGTGAAGCAAGATCGCTA 180

Db      207 gaacacaccttacctgtaaacagagacacactgaaaaggaaggttagtgggaaccgttgc 266
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      181 GAACACACCTTACCTGTAAACAGAGACACACTGAAAAGGAAGGTTAGTGGGAACCGTTGC 240

Db      267 gccagccctgttactggtccaggttcaaagagggatgctcacttctgcgctgtctgcagc 326
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      241 GCCAGCCCTGTTACTGGTCCAGGTTCAAAGAGGGATGCTCACTTCTGCGCTGTCTGCAGC 300

Db      327 gattacgcatcgggatatcactatggagtctggtcgtgtgaaggatgtaaggcctttttt 386
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      301 GATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATGTAAGGCCTTTTTT 360

Db      387 aaaagaagcattcaaggacataatgattatatttgtccagctacaaatcagtgtacaatc 446
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      361 AAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCAGTGTACAATC 420

Db      447 gataaaaaccggcgcaagagctgccaggcctgccgacttcggaagtgttacgaagtggga 506
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      421 GATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTACGAAGTGGGA 480

Db      507 atggtgaagtgtggctcccggagagagagatgtgggtaccgccttgtgcggagacagaga 566
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      481 ATGGTGAAGTGTGGCTCCCGGAGAGAGAGATGTGGGTACCGCCTTGTGCGGAGACAGAGA 540

Db      567 agtgccgacgagcagctgcactgtgccggcaaggccaagagaagtggcggccacgcgccc 626
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      541 AGTGCCGACGAGCAGCTGCACTGTGCCGGCAAGGCCAAGAGAAGTGGCGGCCACGCGCCC 600

Db      627 cgagtgcgggagctgctgctggacgccctgagccccgagcagctagtgctcaccctcctg 686
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      601 CGAGTGCGGGAGCTGCTGCTGGACGCCCTGAGCCCCGAGCAGCTAGTGCTCACCCTCCTG 660

Db      687 gaggctgagccgccccatgtgctgatcagccgccccagtgcgcccttcaccgaggcctcc 746
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      661 GAGGCTGAGCCGCCCCATGTGCTGATCAGCCGCCCCAGTGCGCCCTTCACCGAGGCCTCC 720

Db      747 atgatgatgtccctgaccaagttggccgacaaggagttggtacacatgatcagctgggcc 806
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      721 ATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTACACATGATCAGCTGGGCC 780

Db      807 aagaagattcccggctttgtggagctcagcctgttcgaccaagtgcggctcttggagagc 866
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      781 AAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAAGTGCGGCTCTTGGAGAGC 840

Db      867 tgttggatggaggtgttaatgatggggctgatgtggcgctcaattgaccacccgggcaag 926
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      841 TGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCAATTGACCACCCGGGCAAG 900

Db      927 ctcatctttgctccagatcttgttctggacagggatgaggggaaatgcgtagaaggaatt 986
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      901 CTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGGAAATGCGTAGAAGGAATT 960

Db      987 ctggaaatctttgacatgctcctggcaactacttcaaggtttcgagagttaaaactccaa 1046
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      961 CTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTTCGAGAGTTAAAACTCCAA 1020

Db     1047 cacaaagaatatctctgtgtcaaggccatgatcctgctcaattccagtatgtaccctctg 1106
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     1021 CACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAATTCCAGTATGTACCCTCTG 1080

Db     1107 gtcacagcgacccaggatgctgacagcagccggaagctggctcacttgctgaacgccgtg 1166
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     1081 GTCACAGCGACCCAGGATGCTGACAGCAGCCGGAAGCTGGCTCACTTGCTGAACGCCGTG 1140

Db     1167 accgatgctttggtttgggtgattgccaagagcggcatctcctcccagcagcaatccatg 1226
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     1141 ACCGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATCTCCTCCCAGCAGCAATCCATG 1200

Db     1227 cgcctggctaacctcctgatgctcctgtcccacgtcaggcatgcgag 1273
            |||||||||||||||||||||||||||||||||||||||||||||||
Qy     1201 CGCCTGGCTAACCTCCTGATGCTCCTGTCCCACGTCAGGCATGCGAG 1247


RESULT    2
ID   T62842 standard; cDNA; 2568 BP.
AC   T62842;
DT   08-JUN-1997  (first entry)
DE   Rat oestrogen receptor beta (ER-beta) cDNA.
KW   Orphan receptor; oestrogen receptor beta; ER-beta;
KW   nuclear receptor; prostate cancer; benign prostatic hyperplasia;
KW   ovary cancer; cardiovascular disease; osteoporosis;
KW   environment; pollutant; ss.
OS   Rattus sp.
FH   Key                 Location/Qualifiers
FT   CDS                 424..1881
FT   /*tag=  a
FT   /note= "alternative start codon at 448..450"
PN   WO9709348-A2.
PD   13-MAR-1997.
PF   09-SEP-1996; E03933.
PR   08-SEP-1995; GB-018272.
PR   15-MAR-1996; GB-005550.
PR   11-APR-1996; GB-007532.
PR   08-MAY-1996; GB-009576.
PA   (KARO-) KARO BIO AB.
PI   Enmark E,  Gustafsson J,  Kuiper GG;
DR   WPI; 97-192842/17.
DR   P-PSDB; W14723.
PT   New isolated oestrogen receptor beta   used to develop prods. for
PT   treating e.g. cancers, CNS diseases, osteoporosis or cardiovascular
PT   disease
PS   Claim 6; Fig 1; 45pp; English.
CC   A cDNA clone (T62842) codes for a novel rat oestrogen receptor-
CC   related nuclear receptor (W13723), designated ER-beta.  A set of
CC   degenerate primers based on conserved sequences of the DNA-
CC   binding domain and ligand binding domain of nuclear receptors were
CC   used to amplify rat prostate cDNA.  The PCR prod. was used to
CC   screen a rat prostate cDNA library to obtain the ER beta cDNA.
CC   This was then used to obtain human ER-beta cDNA (see also T62843)
CC   from a human testis cDNA library.  Rat, human and mouse ER-beta
CC   (W14723-25) can be used to isolate molecules for use in the
CC   treatment of cardiovascular diseases, central nervous system
CC   diseases, osteoporosis, prostate or ovarian cancer or benign
CC   prostatic hyperplasia and to test environmental chemicals for
CC   oestrogenic activity.
SQ   Sequence   2568 BP;    634 A;    654 C;    663 G;    617 T;

  Query Match             68.8%;  Score 861;  DB 29;  Length 2568;
  Best Local Similarity  84.7%;  Pred. No. 0.00e+00;
  Matches 1051;  Conservative    0;  Mismatches 190;  Indels    0;  Gaps    0;

Db      448 atgaactacagtgttcccggcagcaccagtaacctggacggtgggcctgtccaactgaac 507
            ||||| |||||  ||||| |||    || |||| |||| ||||||||||  | || | || |
Qy        1 ATGAATTACAGCATTCCCAGCAATGTCACTAACTTGGAAGGTGGGCCTGGTCGGCAGACC 60

Db      508 acaagcccaaatgtgctatggccaacttctgggcacctgtctcctttagcgacccattgc 567
            ||||||||||||||| | ||||||||  |||||||||| |||||||||| |  |||| ||
Qy       61 ACAAGCCCAAATGTGTTGTGGCCAACACCTGGGCACCTTTCTCCTTTAGTGGTCCATCGC 120
```

```
Db        568 caatcatcgctcctctatgcagaacctcaaaagagtccttggtgtgaagcaagatcacta 627
              || | ||| |  || ||||| ||||||||||||||||| ||||||||||||||||| |||
Qy        121 CAGTTATCACATCTGTATGCGGAACCTCAAAAGAGTCCCTGGTGTGAAGCAAGATCGCTA 180

Db        628 gaccacaccttacctgtaaacagagagacactgaaagaggaagcttagtgggagcagttgt 687
              || |||||||||||||||||||||||||||||||||  | || ||| ||||| | |||| 
Qy        181 GAACACACCTTACCTGTAAACAGAGAGACACTGAAAAGGAAGGTTACTGGGAACCGTTGC 240

Db        688 gccagccctgttactagtccaagcaccaagagggatgctcacttctgccctgtctgcagc 747
              ||||||||||||||| |||||| |  | ||  ||||||||||||||||||| ||||||||||
Qy        241 GCCAGCCCTGTTACTGGTCCAGGTTCAAACAGGGATGCTCACTTCTGCGCTGTCTGCAGC 300

Db        748 gattatgcatctgggtatcattacggcgtttggtcatgtgaaggatgtaaggcctttttt 807
              ||||| ||||| || ||||| || || | ||||| |||||||||||||||||||||||||
Qy        301 GATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATGTAAGGCCTTTTTT 360

Db        808 aaaagaagcattcaaggacataatgattatatctgtccagccacgaatcagtgtaccata 867
              |||||||||||||||||||||||||||||||| |||||||| || ||||||||||| || 
Qy        361 AAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCAGTGTACAATC 420

Db        868 gacaagaaccggcgtaaaagctgccaggcctgccgacttcgcaagtgttatgaagtagga 927
              || || |||||||| || |||||||||||||||||||||||| |||||||| ||||| |||
Qy        421 GATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTACGAAGTGGGA 480

Db        928 atggtcaagtgtggatccaggagagaacggtgtgggtacggtatagtgcggaggcagaga 987
              ||||| |||||||| ||| ||||||  | ||||||||| |  | |||||||| ||||||
Qy        481 ATGGTGAAGTGTGGCTCCCGGAGAGAGAGATGTGGGTACCGCCTTGTGCGGAGACAGAGA 540

Db        988 agttctaggaggaggtacactgcctgagcaaagccaagagaaacggtggcatgcaccccc 1047
              |||  |    ||||||| | |||||      |||| ||||||||||  || || || || |||
Qy        541 AGTGCCGACGAGTAGCTGCACTGTGCCGGCAAGGCCAAGAGAAGTGGCGGCCACCCCGCC 600

Db       1048 cgactgaggaactactcctggagcacttgagtccagagcagttggtgcttcaccctcctg 1107
              || |||   ||||| ||||||    | || |||| || ||||| || ||||||||||||||
Qy        601 CGAGTGCGGAGCTGCTGCTGGACCCCTGAGCCCGAGCAGCTAGTGCTCACCCTCCTCCTG 660

Db       1108 gaagctgaaccacccaatgtgctgatgagcagtcccagcatgcccttcacggaggcctgc 1167
              || ||||| || ||| |||||||| | ||||| |||||    ||||||||||||||||||||
Qy        661 GAGGCTGAGCCCCCCATGCGCTGATCAGCCGCCCCAGTGCCCCCCTTCACCGAGGCCTCC 720

Db       1168 atgatgatgtccctcactaagctggcggacaaggaactggtgcacatgattgactgggcc 1227
              |||||||||||||| || ||||||| |||||||| |||||| ||||||||   |||||||
Qy        721 ATGATGATGTCCCTGACCAAGCTGGCCGACAAGGAGCTGGTACACATGATCAGCTGGGCC 780

Db       1228 aagaaattccctgagtttgtggagctcagcctattcgaccaagtccggctcttagagagc 1287
              ||||| |||| | |  |||||||||||||||| |||||||||||| |||||||| ||||||
Qy        781 AAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAAGTGCGGCTCTTGGAGAGC 840

Db       1288 tgctggatggaggtgctaatgatgggactgatgtggcgtcaatcgaccacccccggcaag 1347
              || ||||||||||||| ||||| |||| ||||||||||||||| || ||||||||||||||||
Qy        841 TGTTGGATGGAGGTGTTAATGATGGGCTGATGTGGCGCTCAATTGACCACCCCCGGCAAG 900

Db       1348 ctcatttcgctcccgacctcgttctggacagggatgagggaagtgcgtagaaggagatt 1407
              ||||| || ||||| || || |||||||||||||||||||||||| ||||||||||| |||
Qy        901 CTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGAAATGCGTAGAAGGAATT 960

Db       1408 ctggaaatctttgacatgctcctggcaaccagttcaaggttctgagacttaaaactccag 1467
              |||||||||||||||||||||||||||| || || |||||||| || |||||||||||||||
Qy        961 CTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTTCGAGACTTAAAACTCCAA 1020

Db       1468 cacaaggagtatctctgtatgaaggccatgatcctcctcaattccaatatgtaccccttg 1527
              ||||| || |||||||||| |||||||||||||||| ||||| |||||||||||||  ||
Qy       1021 CACAAAGAATATTTCTGTGTCAAGGCCATGATCCTGCTCAATTCCAGTATGTACCCTCTG 1080

Db       1528 gtttctgccagcacaggagcacaagtagcaggaagctgacacagctactaaacacagatg 1587
              |    | || | |||||| || || || |||||||||| | ||| | |||||||| |||
Qy       1081 GTCACAGCGACCCAGGATGCTGACAGCAGCCGGAAGCTGGCTCACTTGCTGAACGCCGTG 1140

Db       1588 acagatgccctggtctgggtgattgcgaagagtggtatctctctccagcagcagtcagtc 1647
              || |||||   ||| ||||||||||||| ||||| || |||||||||||||||||| ||  |
Qy       1141 ACCGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATCTCCTCCCAGCAGCAATCCATG 1200

Db       1648 cgactggccaacctcctgatgcttctttctcacgtcaggca 1688
```

```
              || |||||| |||||||||||||| || || |||||||||||
Qy       1201 CGCCTGGCTAACCTCCTGATGCTCCTGTCCCACGTCAGGCA 1241

RESULT    3
ID   T62844 standard; cDNA; 1458 BP.
AC   T62844;
DT   08-JUN-1997  (first entry)
DE   Mouse oestrogen receptor beta (ER-beta) cDNA.
KW   Orphan receptor; oestrogen receptor beta; ER-beta;
KW   nuclear receptor; prostate cancer; benign prostatic hyperplasia;
KW   ovary cancer; cardiovascular disease; osteoporosis;
KW   environment; pollutant; ss.
OS   Mus sp.
PN   WO9709348-A2.
PD   13-MAR-1997.
PF   09-SEP-1996; E03933.
PR   08-SEP-1995; GB-018272.
PR   15-MAR-1996; GB-005550.
PR   11-APR-1996; GB-007532.
PR   09-MAY-1996; GB-009576.
PA   (KARO-) KARO BIO AB.
PI   Enmark E,  Gustafsson J,  Kuiper GG;
DR   WPI; 97-192842/17.
DR   P-PSDB; W14725.
PT   New isolated oestrogen receptor beta - used to develop prods. for
PT   treating e.g. cancers, CNS diseases, osteoporosis or cardiovascular
PT   disease
PS   Claim 6; Fig 14B; 45pp; English.
CC   A cDNA clone (T62844) codes for a novel mouse oestrogen receptor-
CC   related nuclear receptor (W13725), designated ER beta.  Rat, human
CC   and mouse ER-beta (W14723-25) can be used to isolate molecules for
CC   use in the treatment of cardiovascular diseases, central nervous
CC   system diseases, osteoporosis, prostate or ovarian cancer or benign
CC   prostatic hyperplasia and to test environmental chemicals for
CC   oestrogenic activity.
SQ   Sequence   1458 BP;    365 A;    373 C;    403 G;    317 T;

  Query Match             67.4%;  Score 843;  DB 29;  Length 1458;
  Best Local Similarity   84.0%;  Pred. No. 0.00e+00;
  Matches 1042;  Conservative    0;  Mismatches 199;  Indels    0;  Gaps    0;

Db         25 atgaactccagtgttcccagcagcacgggtagcctggaaggtgggccttttgcccagact 84
              ||||| ||||| |||||||||    |  |||| |||||||||||||||| ||| |||||
Qy          1 ATGAATTACAGCATTCCTAGCAATGCTACTAACTTGGAAGGTGGGCCTGGCCCCAGACC 60

Db         85 gcaagcccaaatgtgctatggccaacttctggacacctctctcctttagcaacccactgc 144
               |||||||||||||| | ||||||||    |||| ||||| ||||||||||     ||| ||
Qy         61 ACAAGCCCAAATGCGTCCTGGCTAACACTGGGCACCTCCTCCTTTAGCCCCCCACCGC 120

Db        145 caatcatcgctcctctatgcagaacctcaaaagagtccttggtgtgaagcaagatcacta 204
              || | ||| |  || ||||| ||||||||||||||||| ||||||||||||||||| |||
Qy        121 CAGTTATCACATCTGTATGCGGAACCTCAAAAGAGTCCCTGGTGTGAAGCAAGATCGCTA 180

Db        205 gaccacaccttacctgtaaacagagagacactgaaagaggaagcttagtgggagcagttgt 264
              || |||||||||||||||||||||||||||||||||  | || ||| ||||| | |||| 
Qy        181 GAACACACCTTACCTGTAAACAGAGAGACACTGAAAAGGAAGGTTACTGGGAACCGTTGC 240

Db        265 gccagccctgttactagtccaagcaccaagagggatgctcacttctgccctgtctgcagt 324
              ||||||||||||||| |||||| |  | ||  ||||||||||||||||||| ||||||||| 
Qy        241 GCCAGCCCTGTTACTGGTCCAGGTTCAAACAGGGATGCTCACTTCTGCGCTGTCTGCAGC 300

Db        325 gattatgcatctgggtatcattacggcgtttggtcatgtgaaggatgtaaggcctttttt 384
              ||||| ||||| || ||||| || || | ||||| |||||||||||||||||||||||||
Qy        301 GATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATGTAAGGCCTTTTTT 360

Db        385 aaaagaagcattcaaggacataatgattatatctgtccagccacgaatcagtgtaccata 444
              |||||||||||||||||||||||||||||||| |||||||| || ||||||||||| || 
Qy        361 AAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCAGTGTACAATC 420

Db        445 gacaagaaccggcgtaaaagctgccaggcctgccgacttcgcaagtgttatgaagtagga 504
              || || |||||||| || |||||||||||||||||||||||| |||||||| ||||| |||
Qy        421 GATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTACGAAGTGGGA 480
```

```
Db        505 atggtcaagtgtggatccaggagagaaaggtgtgggtaccgaatagtacgaagacagaga 564
              ||||| |||||||| ||| ||||||| || |||||||||||  | || || |||||||||
Qy        481 ATGGTGAAGTGTGGCTCCCGGAGAGAGAGATGTGGGTACCGCCTTGTGCGGAGACAGAGA 540

Db        565 agtgccagcgagcaggtgcattgcctgaacaaagccaagagaagtggccagcacacaccc 624
              ||||||  ||||||| |||| ||      ||| |||||||||||||||   ||| | |||
Qy        541 AGTGCCGACGAGCAGCTGCACTGTGCCGGCAAGGCCAAGAGAAGTGGCGGCCACGCGCCC 600

Db        625 cgaatgaaggagctactgctgaactctctgagtcccgagcagctggtgctcaccctgctg 684
              ||| ||  |||||| |||||| || | ||||| ||||||||||| ||||||||||| |||
Qy        601 CGAGTGCGGGAGCTGCTGCTGGACGCCCTGAGCCCCGAGCAGCTAGTGCTCACCCTCCTG 660

Db        685 gaagctgagccacccaatgtgctagtgagtcgtcccagcatgcccttcaccgaggcctcc 744
              || |||||||| ||| |||||||  | || || |||||   |||||||||||||||||||
Qy        661 GAGGCTGAGCCGCCCCATGTGCTGATCAGCCGCCCCAGTGCGCCCTTCACCGAGGCCTCC 720

Db        745 atgatgatgtcccttacgaagctggctgacaaggaactggtgcacatgattggctgggcc 804
              |||||||||||||| || ||| |||| ||||||||  |||| ||||||||  ||||||||
Qy        721 ATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTACACATGATCAGCTGGGCC 780

Db        805 aagaaaatccctggctttgtggagctcagcctgttggaccaagtccgcctcttggaaagc 864
              ||||| || || ||||||||||||||||||||||| |||||||| || |||||||| |||
Qy        781 AAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAAGTGCGGCTCTTGGAGAGC 840

Db        865 tgctggatggaggtgctgatgatgggactgatgtggcgctccatcgaccaccccggcaag 924
              || |||||||||||| | |||||||| |||||||||||||| || |||||||||||||||
Qy        841 TGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCAATTGACCACCCCGGCAAG 900

Db        925 ctcatctttgctccagacctcgttctggacagggatgaggggaagtgcgtggaagggatt 984
              ||||||||||||||||| || ||||||||||||||||||||||| ||||| ||||| |||
Qy        901 CTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGGAAATGCGTAGAAGGAATT 960

Db        985 ctggaaatctttgacatgctcctggcgacgacggcacggttccgtgagttaaaactgcag 1044
              |||||||||||||||||||||||||| || ||  || |||| || ||||||||||| || 
Qy        961 CTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTTCGAGAGTTAAAACTCCAA 1020

Db       1045 cacaaagaatatctgtgtgtgaaggccatgattctcctcaactccagtatgtaccacttg 1104
              |||||||||||||| ||||| ||||||||||| || ||||| |||||||||||||   ||
Qy       1021 CACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAATTCCAGTATGTACCCTCTG 1080

Db       1105 gctacagcgaagcaggaagcagagagtagccggaagctgacacacctattgaacgcagtg 1164
              |  |||||||  ||||| || || || |||||||||||| | ||| |  ||||||| |||
Qy       1081 GTCACAGCGACCCAGGATGCTGACAGCAGCCGGAAGCTGGCTCACTTGCTGAACGCCGTG 1140

Db       1165 acagatgccctggtctgggtgatttcgaagagtagaatctcttcccagcagcagtcagtc 1224
              || |||||  |||| ||||||||| | |||||  | ||||| ||||||||||| ||| | 
Qy       1141 ACCGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATCTCCTCCCAGCAGCAATCCATG 1200

Db       1225 cgtctagccaacctcctgatgcttctttctcatgtcaggca 1265
              || || || |||||||||||||| || || || ||||||||
Qy       1201 CGCCTGGCTAACCTCCTGATGCTCCTGTCCCACGTCAGGCA 1241


RESULT    4
ID   N70880 standard; cDNA; 2092 BP.
AC   N70880;
DT   24-FEB-1991  (first entry)
DE   cDNA encoding human oestrogen receptor protein
KW   Human oestrogen receptor protein; ss cDNA, expression system, assay,
KW   steroid receptor protein;
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   CDS             293..1280
FT   /*tag   a
PN   WO8705049-A.
PD   27 AUG-1987.
PF   18-FEB-1987; U00341.
PR   20-FEB-1986; US-833829.
PA   (CALB-) California Biotechnology Inc
PI   Shine J;
DR   WPI; 87-250213/35
DR   P-PSDB; P70543.
PT   Expression system for vertebrate steroid receptor protein - comprising
PT   DNA sequence encoding the protein linked to control sequences in
PT   eucaryotic hosts.
PS   Disclosure; Fig. 1-1 - 1-2; 26pp; English.
CC   The cDNA can be used in an expression system to express human
CC   oestrogen receptor protein. The coding sequence is operably linked to
CC   control sequences compatible with eukaryotic host cells. This method
CC   allows expression under conditions which favour appropriate post-
CC   translational processing. It produces large amts. of purified protein
CC   useful in the design of agonist and antagonist cpds., for the study of
CC   the mechanism of action of the steroid binding proteins in general,
CC   and for use in diagnostic assays for the proteins or antibodies to
CC   them. These assays are important in, eg the diagnosis of tumour
CC   sensitivities to steroid metabolism. Suitable host cells are VERO,
CC   HeLa and CHO cells.
SQ   Sequence   2092 BP;   [illegible] A;   [illegible] C;   591 G;   421 T;

  Query Match             10.9%;  Score 135;  DB 2;  Length 2092;
  Best Local Similarity   77.8%;  Pred. No. 2.37e-76;
  Matches  189;  Conservative    0;  Mismatches   54;  Indels    0;  Gaps    0;

Db        832 ggagactcgctactgtgcagtgtgcaatgactatgcttcaggctaccattatggagtctg 891
              |||  ||| || |||  | || ||||  || || || || || || |||||||||||
Qy        273 GGATGCTCACTTCTGCCCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTG 332

Db        892 gtcctgtgagggctgcaaggccttcttcaagagaagtattcaaggacataacgactatat 951
              ||| ||||| || || |||||||| || || ||||| |||||||||||||| || |||||
Qy        333 GTCGTGTGAAGGATGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATAT 392

Db        952 gtgtccagccaccaaccagtgcaccattgataaaaacaggaggaagagctgccaggcctg 1011
               |||||||| || || ||||| || || ||||||||| || | |||||||||||||||||
Qy        393 TTGTCCAGCTACAAATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTG 452

Db       1012 ccggctccgcaaatgctacgaagtgggaatgatgaaaggtgggatacgaaaagaccgaag 1071
              ||| || || || || ||||||||||||||| ||||  ||||    || | |||  || |
Qy        453 CCGACTTCGGAAGTGTTACGAAGTGGGAATGGTGAAGTGTGGCTCCCGGAGAGAGAGATG 512

Db       1072 agg 1074
               ||
Qy        513 TGG 515


RESULT    5
ID   T06873 standard; DNA; 4963 BP.
AC   T06873;
DT   13-MAR-1996  (first entry)
DE   PhCMV*-1 promoter and human oestrogen receptor.
KW   Transactivator, tTA, tet operator; teto; gene expression;
KW   tetracycline-responsive promoter; PhCMV*-1; human cytomegalovirus;
KW   transcription activation; eukaryotic cell; oestrogen receptor;
KW   ds; ss.
OS   Chimeric synthetic;
OS   Chimeric Homo sapiens.
PN   US5464758-A.
PD   07-NOV-1995.
PF   14-JUN-1993; 076726.
PR   14-JUN-1993; US-076726.
PA   (BUJA/) BUJARD H.
PA   (GOSS/) GOSSEN M.
PI   Bujard H,  Gossen M;
DR   WPI; 95-392612/50.
PT   Polynucleotide encoding trans-activator fusion protein contg. tet
PT   repressor - used to control expression of gene regulated by minimal
PT   promoter linked to tet operon, and vectors and cells where gene
PT   expression is regulated by tetracycline
PS   Disclosure; Fig 10A-B; 37pp; English.
CC   cDNA coding for the human oestrogen receptor was placed under
CC   control of the PhCMV*-1 promoter (see T06869), resulting in
CC   tetracycline-regulated expression vector pUHDpgr-3 contg. the
CC   sequence given in T06873. The construct was transfected into
CC   eukaryotic cells expressing the tTA transactivator (T06867) to
CC   allow tetracycline-regulated expression of the receptor.
SQ   Sequence   4963 BP;    1259 A;    1300 C;    1278 G;    1126 T;
```

```
  Query Match             10.8%;  Score 135;  DB 17;  Length 4963;
  Best Local Similarity   77.8%;  Pred. No. 2.37e-76;
  Matches  189;  Conservative    0;  Mismatches   54;  Indels    0;  Gaps    0;

Db   1005 ggagactcgctactgtgcagtgtgcaatgactatgcttcaggctaccattatggagtctg 1064
          |||  ||| || ||| || || ||||  || || || || || || || |||||||||||
Qy    273 GGATGCTCACTTCTGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTG 332

Db   1065 gtcctgtgagggctgcaaggccttcttcaagagaagtattcaaggacataacgactatat 1124
          ||| ||||| || || |||||||| || || ||||| |||||||||||||| || |||||
Qy    333 GTCGTGTGAAGGATGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATAT 392

Db   1125 gtgtccagccaccaaccagtgcaccattgataaaaacaggaggaagagctgccaggcctg 1184
           |||||||| || || ||||| || || ||||||||| || | |||||||||||||||||
Qy    393 TTGTCCAGCTACAAATCAGTGTACAATCGATAAAAACTGGCGCAAGAGCTGCCAGGCCTG 452

Db   1185 ccggctccgcaaatgctacgaagtgggaatgatgaaaggtgggatacgaaaagaccgaag 1244
          ||| || || || || ||||||||||||||| ||||  ||||    || | |||| || |
Qy    453 CCGACTTCGGAAGTGTTACGAAGTGGGAATGGTGAAGTGTGGCTCCCGGAGAGACAGATG 512

Db   1245 agg 1247
           ||
Qy    513 TGG 515


RESULT    6
ID   Q76270 standard; DNA; 4963 BP.
AC   Q76270;
DT   17-JUL-1995  (first entry)
DE   PhCMV*-1-controlled progesterone receptor
KW   tTA; transactivator; tetracycline-controllable transactivator;
KW   conditional inactivation; homologous recombination; gene expression;
KW   gene regulation; gene therapy; tetracycline-resistance; tetR
KW   transgenic animal; PhCMV*-1; promoter; tetO; CMV;
KW   progesterone receptor; ds.
OS   Human cytomegalovirus Towne.
PN   WO9429442-A.
PD   22-DEC-1994.
PF   14-JUN-1994; U06734.
PR   14-JUN-1993; US-076327.
PA   (BADI ) BASF AG.
PI   Bujard H,  Gossen M,  Salfeld JG,  Voss JW;
DR   WPI; 95-026472/05.
PT   Regulatory systems using tetracycline-controllable transactivator
PT   (tTA) - useful for conditional inactivation or modulation of
PT   gene expression in a host cell or animal
PS   Disclosure; Page 54-58; 103pp; English.
CC   The minimal promoter PhCMV*-1 is based on the human CMV
CC   intermediate-early promoter (with enhancer region removed), fused
CC   to a tetracycline operator, tetO.  The promoter is used to
CC   express tTA transactivator in host cells.  In clone pUHD 80F4
CC   (given in Q76269), cDNA for rabbit progesterone receptor is under
CC   the control of PhCMV*-1.
SQ   Sequence   4963 BP;   1259 A;   1298 C;   1280 G;   1126 T;

  Query Match             10.8%;  Score 135;  DB 13;  Length 4963;
  Best Local Similarity   77.8%;  Pred. No. 2.37e-76;
  Matches  189;  Conservative    0;  Mismatches   54;  Indels    0;  Gaps    0;

Db   1005 ggagactcgctactgtgcagtgtgcaatgactatgcttcaggctaccattatggagtctg 1064
          |||  ||| || ||| || || ||||  || || || || || || || |||||||||||
Qy    273 GGATGCTCACTTCTGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTG 332

Db   1065 gtcctgtgagggctgcaaggccttcttcaagagaagtattcaaggacataacgactatat 1124
          ||| ||||| || || |||||||| || || ||||| |||||||||||||| || |||||
Qy    333 GTCGTGTGAAGGATGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATAT 392

Db   1125 gtgtccagccaccaaccagtgcaccattgataaaaacaggaggaagagctgccaggcctg 1184
           |||||||| || || ||||| || || ||||||||| || | |||||||||||||||||
Qy    393 TTGTCCAGCTACAAATCAGTGTACAATCGATAAAAACTGGCGCAAGAGCTGCCAGGCCTG 452

Db   1185 ccggctccgcaaatgctacgaagtgggaatgatgaaaggtgggatacgaaaagaccgaag 1244
          ||| || || || || ||||||||||||||| ||||  ||||    || | |||| || |
Qy    453 CCGACTTCGGAAGTGTTACGAAGTGGGAATGGTGAAGTGTGGCTCCCGGAGAGACAGATG 512

Db   1245 agg 1247
           ||
Qy    513 TGG 515


RESULT    7
ID   Q85976 standard; cDNA; 204 BP.
AC   Q85976;
DT   31-OCT-1995  (first entry)
DE   Human estrogen receptor cDNA.
KW   nuclear receptor protein; osteosarcoma; EAR-1r; generation; growth;
KW   differentiation; human; erb-A related gene; estrogen receptor; ds.
OS   Homo sapiens.
PN   J07033798-A.
PD   03-FEB-1995.
PF   20-JUL-1993; 200333.
PR   20-JUL-1993; JP-200333.
PA   (MITU ) MITSUBISHI KASEI CORP.
DR   WPI; 95-110650/15.
DR   P-PSDB; R71566.
PT   Nuclear receptor protein derived from osteo blastic osteo-sarcoma
PT   - encoded by the EAR-1r gene, for investigation of generation of
PT   fat soluble active substances
PS   Disclosure; Page 6; 8pp; Japanese.
CC   This cDNA encodes an estrogen receptor (a member of a receptor
CC   superfamily). It is disclosed in the patent specification. The
CC   invention concerns an EAR 1r gene (human erb-A related gene)
CC   encoding a nuclear receptor protein (R71565) derived from human
CC   osteosarcoma (HOS) cells. The expression of EAR-1r is most abundant
CC   in the brain and was detectedd slightly in heart and placenta. The
CC   receptor protein may be used for the investigation of generation,
CC   differentiation and growth of fat soluble physiological substances.
SQ   Sequence   204 BP;   59 A;   46 C;   55 G;   44 T;

  Query Match             10.2%;  Score 127;  DB 14;  Length 204;
  Best Local Similarity   81.3%;  Pred. No. 1.64e-70;
  Matches  165;  Conservative    0;  Mismatches   38;  Indels    0;  Gaps    0;

Db      1 tgtgcagtgtgcaatgactatgcttcaggctaccattatggagtctggtcctgtgagggc 60
          || || || ||||  || || || || || || || |||||||||||||| ||||| ||
Qy    286 TGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGA 345

Db     61 tgcaaggccttcttcaagagaagtattcaaggacataacgactatatgtgtccagccacc 120
          || |||||||| || || ||||| |||||||||||||| || ||||| |||||||| ||
Qy    346 TGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACA 405

Db    121 aaccagtgcaccattgataaaaacaggaggaagagctgccaggcctgccggctccgcaaa 180
          || ||||| || || ||||||||| || | ||||||||||||||||||||| || || ||
Qy    406 AATCAGTGTACAATCGATAAAAACTGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAG 465

Db    181 tgctacgaagtgggaatgatgaa 203
          || ||||||||||||||| ||||
Qy    466 TGTTACGAAGTGGGAATGGTGAA 488


RESULT    8
ID   Q10572 standard; DNA; 1047 BP.
AC   Q10572;
DT   09-APR-1991  (first entry)
DE   Human Natriuretic Peptide Receptor B.
KW   NPRB; ANP; BNP; CNP; kidney failure; heart failure; protein kinase;
KW   hyperaldosteronism; glaucoma; guanyl cyclase.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   Peptide         1..22
FT   /label= signal sequence
FT   Protein         12
FT   /label= mature NPBR
FT   Domain          23..455
FT   /label= extracellular domain
FT   /note= "binds natriuretic peptides A,B and C]"
FT   Domain              456..456
FT   /label= transmembrane domain
FT   Domain              479..1047
FT   /label= cytoplasmic domain
FT   /note= "GC and protein kinase activity"
FT   Modified -site  24..26
FT   /label= N-glycos_site
FT   Modified -site  35..37
FT   /label= N-glycos_site
FT   Modified -site  161..163
FT   /label= N-glycos_site
FT   Modified -site  195..197
FT   /label= N-glycos_site
FT   Modified -site  244..246
FT   /label= N-glycos_site
FT   Modified -site  277..279
FT   /label= N-glycos_site
FT   Modified -site  349..351
FT   /label= N-glycos_site
FT   Modified -site  600..602
FT   /label= N-glycos_site
PN   WO9100292-A.
PD   10-JAN-1991.
PF   22-JUN-1990; U03586.
PR   23-JUN-1989; US 370673.
PA   (GETH ) GENENTECH INC.
PI   Chang M, Goeddel D, Lowe D;
DR   WPI; 91-036711/05.
DR   N-PSDB; Q10324
PT   Natriuretic protein receptor B  for diagnosis and treatment of
PT   kidney failure, heart failure, hyperaldosteronism, glaucoma etc.
PS   Claim 3; Fig 1; 49pp; English.
CC   The sequence was derived from the DNA encoding natriuretic peptide
CC   receptor B, NPRB, having guanyl cyclase (GC) activity and protein
CC   kinase activity.  The DNA can be inserted into expression vectors
CC   for the prodn. of the protein, opt. after being mutated to produce
CC   NPRB analogues. The protein has a mol wt. of 115 kD (calculated Mr=
CC   114,952).   The protein (or variants) can be used in treatment of
CC   natriuretic peptide disorders, and also to isolate peptides using
CC   affinity chromatography.  Antibodies with affinity for NPRB can
CC   also be prepd.
SQ   Sequence   1047 BP;    87 A;    15 C;    83 G;    51 T;

  Query Match              7.3%;  Score 91,  DB 2,  Length 1047,
  Best Local Similarity    9.9%;  Pred. No. 1.24e-44,
  Matches    83;  Conservative  269,  Mismatches 572,  Indels  10,  Gaps  10,

Db      108 varnashwrnnnntaqavasqnsakndhyrtnvrtqnsankngnnvvtnhghnnwtaraa 167
               :|: |:: :          ||:| |   !: :. .| ..   .   .    :.  .  .  .
Qy      198 AAACAGAGAGACACTGAAAAGCAAGGTTAGTGGGAACCGTTGCGCCAGCCCTGTTACTGG 257

Db      168 nnyndart-ddrnhyntnnqvnnannqsnnsvnhnvyarnnqgnnnathnnrangrnvyn 226
                : :  | ::: :       |:   |     :   :: :   :  :   |         :   :   |: ::
Qy      258 TCCAGGTTCAAAGAGGGATGCTCACTTCTGCGCTGTCTGCAGCGATTACCCATCGGGATA 317

Db      227 qrrnmnhnnnnnarrnnntn-adyvnnyndvnqncnraqntratgrnwndnrtrnnaaa 285
                      :          :     |   |: :   : :: | :       | |   :  :  :  :|    |
Qy      318 TCACTATGGAGTCTGGTCGTGTGAAGGATGTAAGGCCTTTTTAAAAGAAGCATTCAAGG 377

Db      286 nrnannnvnvntyrnnnnnnynnnnrnnnrarndnqvnnqnsnmnnnagcnydgnnnya 345
                          |:   ::       :                  : :  :    : :    |   ::|    |
Qy      378 ACATAATGATTATATTTGTCCAGCTACAAATCAGTGTACAATCGATAAAAACGGGCGCAA 437

Db      346 nvnnntnnnqtrndqnrnvnkmnqrryhqvtqnvvmdknndrntdnvnwamqdndsgdn 405
                  :       ||     :   :  :  ::  |::  :| |    ::::   :: |:     :|   :  :: :
Qy      438 GAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTACGAAGTGGGAATGGTGAAGTGTGGCTC 497

Db      406 nnaahysqanknnwwtqrnnnwvkqannsdnnncandnddnsedktnnstnanvangtqn 465
                   :   ||     : ||:       ::        ::       | : :: : ::|   :   | :|    |
Qy      498 CCGGAGAGAGATGTGGGTACCGCCTTGTGCGGAGACAGAACTGCC-CACCACCAGC 556

Db      466 tnnmnqvssnnnnrknmnnknnasnnwrnrwnnnnnnnsnrykkqagsrntnsnrqssyg 525
           |  :  :       :  :        |:: :: ::             ::        :: |  :  :| ::
Qy      557 TGCACTGTGCCG-GCAAGGCCAAGAGAA-GTGGCGGCCACGCGCCCCGAGTGCGGGAGCT 614

Db      526 snmtahgkynnnantqhnkgnvvankhvnkkrnnntrnvnnnnkhmrdvnnnhntrnnqa 585
            :     :|:        ||:       ::|   ::  :  :      |   :          :  :::
Qy      615 GCTGCTTGACGCCCTGACCCCCAGCAGCTACCGCTCACCCTCCTGCACCCCCACCCCCC 674

Db      586 cndnnnncnvtnycntgsnndnnndsnnndwmnrysnnndnvkqmannhnsnnsshqsn 645
            | :      |  :   :     :          ::        :   :::      :  ::  :|    :  :    :  :  :
Qy      675 CCATGTGCTGATCAGCCGCCCCAGTGCGCCCTTCACCGAGGCCTCCATGATGATGTCCCT 734

Db      646 kssncvvdsrnvnkntdygnasn-rstannddnnanyakknntannnnsgnnnnttmnna 704
            :  :   :::  :   :   :     :  | |:     ::||      :     | :|:            :          |    :
Qy      735 GACCAAGTTGGCCGACAAGGAGTTGGTACACATGATCAGCTGGGCCAAGAAGATTCCCGG 794

Db      705 advysnannnnnnanrsgnnynngndnsnknnvnkvrnqnrnynrnsndrtnnnnnnvnn 764
              :  ::  |          | ::      :    |         :     :  ::       :       :   :   :  |          :
Qy      795 CTTTGTGGAGCTCA-GCCTGTTCGACCAAGTGCGGCTCTTGGAGAGCTGTTGGATGGAGG 853

Db      765 mnrcwandnanrndnqnnkgnnrrnnknqqtsnndnnnnrmnnyannnnknvnnrtnayn 824
                  :|  :     :  :       :|   :      :       |                   :    :|                 |   :
Qy      854 TGTTAATGATGGGGCTGATGTGGTGCTCAATTGACCACCCCCGCAAGCTCATCTTTGCTC 913

Db      825 nnkrkanannynnnnhsvannnkrgntvnanandsvtnynsdnvgntans-anstnmnvv 883
              . . .       :        . . .        . .        .  |  |  : : :|        :  :    |   :   |    |        :
Qy      914 CAGATCTTGTTCTGGACAGGGATGAGGGGAAATGCGTAGAAGGAATTCTGGAAATCTTTG 973

Db      884 tnnndnytcndanndnndvykvntngdaymvvsqnnqrnqnrhannnarmananndavss 943
                 :  :  |   :       :    :::::        |: :::::          :          :           ::|  ;   :   :   :
Qy      974 ACATGCTCCTGGCAACTACTTCAAGGTTTCGACAGTTAAAACTCCAACACAAAGAATATC 1033

Db      944 nrnrhrnhdrnrnrnqnhtgnvnagvvgnkrnrycnnqdtnrtnvrrnsnqnarknlnk 1002
                :::  ::    :    :   |::       :|    ::          :  ::         :   :   |    :      :           :  :::
Qy     1034 TCTGTGTCAAGGCCATGATCCTGCTCAATTCCAGTATGTACCCTCTGGTCACAGCGACCC 1093

Db     1003 -sttkdandnngcnnnnnrgdvnmkgkgkmrtyw 1035
               :   ::    :                 :|:: ::|: ::: ::
Qy     1094 ACCATGCTGACAGCAGCCGGAAGCTGGCTCACTT 1127


RESULT    9
ID   T42193 standard; DNA; 3460 BP.
AC   T42193;
DT   20-JAN-1997  (first entry)
DE   Fragment of pig oestrogen receptor gene.
KW   Pig, oestrogen receptor, exon, polymorphism, allele, RFLP, litter size;
KW   restriction fragment length polymorphism, Chinese, American; PCR; primer;
KW   amplification; polymerase chain reaction; ds.
OS   Sus scrofa.
FH   Key               Location/Qualifiers
FT   misc_feature      303..304
FT   /*tag=  a
FT   /note= "undetermined length gap missing in DNA sequence
FT   between these nucleotides"
FT   exon              534..600
FT   /*tag=  b
FT   /note= "exon 3"
FT   misc_feature      1197..1198
FT   /*tag=  c
FT   /note= "approx. 100 200 bases missing from DNA sequence
FT   between these nucleotides"
FT   primer_bind       1898..1915
FT   /*tag=  d
FT   /note= "binds primer A10RII (T42200)"
FT   primer_bind       2011..2031
FT   /*tag=  e
FT   /note= "binds primer DV028 (T42198)"
FT   primer_bind       2146..2171
FT   /*tag=  f
FT   /note= "binds primer ESR REV (T42195)"
FT   primer_bind       2636
FT   /*tag=  g
FT   primer_bind       2636..2663
FT   /*tag= h
FT   /note= "binds primer ESRSR (T42197)"
FT   variation        2688..2693
FT   /*tag= i
FT   /note= "polymorphic PvuII site"
FT   primer_bind      2729..2755
FT   /*tag= j
FT   /note= "binds primer ESRSF (T42196)"
FT   primer_bind      2922..2940
FT   /*tag= k
FT   /note= "binds primer P1F (T42201)"
FT   primer_bind      2992..3002
FT   /*tag= l
FT   /note= "binds primer DV029 (T42199)"
FT   primer_bind      3331..3355
FT   /*tag= m
FT   /note= "binds primer ESR FOR (T42194)"
FT   misc_feature     3123..3128
FT   /*tag= n
FT   /note= "constant PvuII site"
PN   US5550024-A.
PD   27-AUG-1996.
PF   19-APR-1991; 687708.
PR   19-APR-1991; US-687708.
PR   19-OCT-1992; US-961819.
PR   26-SEP-1994; US-312312.
PA   (BIOT-) BIOTECHNOLOGY RES & DEV CORP.
PA   (IOWA ) UNIV IOWA STATE RES FOUND INC.
PI   Jacobson CD, Mileham AJ, Plastow GS, Rothschild MF,
PI   Tuggle CK, Vaske DA;
DR   WPI; 96-401592/40.
PT   Screening pigs for high litter size - by identifying a polymorphism
PT   in the pig oestrogen receptor gene in a DNA sample
PS   Disclosure; Column 19-24; 26pp; English.
CC   This is a fragment of the pig oestrogen receptor (ESR) gene including
CC   exon 3.  The sequence covers the region contg. a polymorphic PvuII
CC   restriction enzyme site which is used as site for determining the ESR
CC   allele carried by the animal.  Restriction fragment length polymorphism
CC   (RFLP) analysis of this region revealed two alleles of 4.3 and 3.7 kb.
CC   Both alleles are involved with increased pig litter size dependent on the
CC   breed of pig involved e.g. Chinese or American pigs.  However pigs with
CC   the 3.7 kb allele had between 1.78-2.4 more pigs born in each litter and
CC   1.8-2.4 more pigs born alive than pigs with the 4.3 kb allele. The allele
CC   can also be determined by PCR using the primers (T42194-T42201).
SQ   Sequence   3460 BP;    1052 A;    765 C;    634 G;    1009 T;

  Query Match              5.7%;  Score 71;  DB 24;  Length 3460;
  Best Local Similarity   83.3%;  Pred. No. 1.13e-30;
  Matches    95;  Conservative    0;  Mismatches   18;  Indels    1;  Gaps    1;

Db     490 ttcatcatgcccacttcgtagcacttggtagccggcaggc-tgacagctcttcctccta 548
           |||| ||| |||||||||||| ||||| || || |||||||| || ||||||| | || |
Qy     488 TTCACCATTCCCACTTCGTAACACTTCCGAAGTCGGCAGGCCTGGCAGCTCTTCCGCCGC 429

Db     549 ttcttatcaattgtgcactggttggtagctggacacatgtagtcattatgtcct 602
           || ||||| ||||| |||| || |||||||||||| || || |||||||||||
Qy     428 TTTTTATCGATTGTACACTGATTTGTAGTTGCACAAATATAATCATTATGTCCT 375


RESULT   10
ID   N80922 standard; cDNA; 2153 BP.
AC   N80922;
DT   29-DEC-1990  (first entry)
DE   Sequence encoding human estrogen-related receptor protein (hERR2)
KW   Hormone receptor; hormone-binding; transcription activation; ss.
OS   Homo sapiens.
FH   Key              Location/Qualifiers
FT   CDS              100..1401
FT   /*tag= a
PN   WO8803168-A.
PD   05-MAY-1988.
PF   23-OCT-1987; U02782.
PR   4-OCT-1986; US-922585
```

```
PR   20-OCT-1987; US-108471.
PA   (SALK) Salk Inst for Biol Stud.
PI   Evans RM, Weinberger CA, Hollenberg SM, Giguere V;
DR   WPI; 88-133242/19.
DR   P-PSDB; P80931.
PT   Recombinant DNA encoding hormone recpetors -
PT   comprising glucocorticoid, mineralocorticoid, thyroid hormone
PT   and novel hormone receptors
PS   Claim 16; Fig V-2(B)-1 and -2; 243pp; English.
CC   DNAs encoding hormone receptors and the hormone receptors themselves are
CC   claimed. The DNA can be used to make the hormone receptor proteins and
CC   functional modified forms in quantities not previously possible. The
CC   receptor proteins can be used to screen cpds. for receptor-agonist or
CC   receptor-antagonist activity. They can also be used in diagnostic assays.
CC   Also claimed is pure DNA encoding oestrogen-related receptors hERR1 and
CC   hERR2 and hERR1 and hERR2. The new hERR1 and hERR2 receptors will provide
CC   the basis for development of an assay system that will lead to the
CC   identification of novel hormones.
SQ   Sequence   2153 BP;    445 A;    627 C;    624 G;    457 T;

  Query Match              5.4%;  Score 68;  DB 2;  Length 2153;
  Best Local Similarity   67.2%;  Pred. No. 1.28e-28;
  Matches   133;  Conservative    0;  Mismatches   65;  Indels    0;  Gaps    0;

Db     412 gtgtgcggggacattgcttctggctaccactatggagtggcctcctgcgaggcttgcaag 471
           || ||| | ||     || || || || ||||||||||     || || ||  |  || |||
Qy     292 GTCTCCATCCATTATCCATCCCGATATCATTATGGAGTCTGGTCGTGTGAACCACCCAAG 351

Db     472 gcgttcttcaagaggaccattcaaggaaacattgaatacagctgcccttgccacccaacgac 531
           || || || || |||| ||||||||| | |   || || |   || || || || ||   ||
Qy     352 GTTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTCTCCAGCTACAAATCAG 411

Db     532 tgtgagattaccaaacgggaggcacaagctcctgctcaggcctgccaggttcatgaagtgccta 591
           |||   |||   |||    |||||||| ||| ||||||||||| |    | | ||    |
Qy     412 TGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTAC 471

Db     592 aaagtggggatgctgaag 609
            ||||||| ||| |||||
Qy     472 GAAGTGGGAATGGTGAAG 489


RESULT   11
ID   Q10572 standard; DNA; 1047 BP.
AC   Q10572;
DT   09-APR-1991  (first entry)
DE   Human Natriuretic Peptide Receptor B.
KW   NPRB, ANF, BNP, CNP, kidney failure, heart failure, protein kinase;
KW   hyperaldosteronism, glaucoma, guanyl cyclase.
OS   Homo sapiens.
FH   Key              Location/Qualifiers
FT   Peptide          1..22
FT   /label= signal sequence
FT   Protein          12
FT   /label= mature NPBR
FT   Domain           23..455
FT   /label= extracellular domain
FT   /note= "binds natriuretic peptides A,B and C}"
FT   Domain           456..456
FT   /label= transmembrane domain
FT   Domain           479..1047
FT   /label= cytoplasmic domain
FT   /note= "GC and protien kinase activity"
FT   Modified -site  24..26
FT   /label= N-glycos_site
FT   Modified -site  35..37
FT   /label= N-glycos_site
FT   Modified -site  161..163
FT   /label= N-glycos_site
FT   Modified -site  195..197
FT   /label= N-glycos_site
FT   Modified -site  244..246
FT   /label= N-glycos_site
FT   Modified -site  277..279
FT   /label= N-glycos_site
FT   Modified -site  349..351
FT   /label= N-glycos_site
FT   Modified -site  600..602
FT   /label= N-glycos_site
PN   WO9100292-A.
PD   10-JAN-1991.
PF   22 JUN-1990; U03586.
PR   23-JUN-1989; US-370673.
PA   (GETH ) GENENTECH INC.
PI   Chang M, Goeddel D, Lowe D;
DR   WPI; 91-036711/05.
DR   N-PSDB; Q10324.
PT   Natriuretic protein receptor B - for diagnosis and treatment of
PT   kidney failure, heart failure, hyperaldosteronism, glaucoma etc.
PS   Claim 3; Fig 1; 49pp; English.
CC   The sequence was derived from the DNA encoding natriuretic peptide
CC   receptor B, NPRB, having guanyl cyclase (GC) activity and protein
CC   kinase activity.  The DNA can be inserted into expression vectors
CC   for the prodn. of the protein, opt. after being mutated to produce
CC   NPRB analogues. The protein has a mol wt. of 115 kD (calculated Mr=
CC   114,952).   The protein (or variants) can be used in treatment of
CC   natriuretic peptide disorders, and also to isolate peptides using
CC   affinity chromatography.  Antibodies with affinity for NPRB can
CC   also be prepd.
SQ   Sequence   1047 BP;    87 A;    15 C;    83 G;    51 T;

  Query Match             5.3%;  Score 66;  DB 2;  Length 1047;
  Best Local Similarity   9.2%;  Pred. No. 2.95e-27;
  Matches   93;  Conservative  268;  Mismatches 646;  Indels   9;  Gaps    9;

Db      18 rnngarnntnavvnnnhnnsyawawnrvqnavanavnangrannvdnrnvssnnngacsn 77
              :  |  :   | |::    .     .         .   | . |  | .          .   .  . .
Cp    1068 ACTGGAATTGAGCAGGATCATGGCTTGACACAGAGATATTCTTTGTGTTGGAGTTTTAA 1009

Db      78 ynannsavdnknylhdnnnnqnqvvynaasvarnashwrnnnntagavasgnsakndhyr 137
              :      :|::    :  :  :        ||:     |::|     ::..      |   |.|     .   .   .
Cp    1008 CTCTCGAAACCTTCAAGTAGTTGCCAGGAGCATGTCAAAGATTTCCAGAATTCCTTCTAC 949

Db     138 tnvrtqnsanknqnnvvtnhqhnnwtaraannyndartddrnhyntnnqvnnanngsnns 197
                :  |   :    :       ::|  :   :    :   |  ||   :   :   :   :   :   :     |    |:          |:
Cp     948 GCATTTCCCCTCATCCCTGTCCAGA-ACAAGATCTGGAGCAAAGATGAGCTTGCCGGGGT 890

Db     198 vnhnvyarnnqnnnathnnrangrnvyncqnnnmnhnnnnanrnnntnqdyvnnyndv 257
              :  :  :         |         |  :    :           ::  |       :   :         |            |  :.        .
Cp     889 GGTCAATTGAGCGCCACATCAGCCCCAT-CATTAACACCTCCATCCAACAGCTCTCCAAG 831

Db     258 nqnsnraqntratqrnwdnr-tnnnananrnanntvnvntyrnnnnnnynnnnnrnnnr 316
               |  :    |    |:  |  :   :   :   :   |:            :  |      :    :      :            :         :     :
Cp     830 AGCCGCACTTGGTCGAACAGCTGAGCTCCACAAAGCCGGGAATCTTCTTGGCCCAGCTG 771

Db     317 arndnqvnnqnsnmnnnaqcnydqnnnyanvnnntnnnqqtrndqnrnvnkmngtryhgv 376
              |   :  |         :   :              :  |        |          |       ||       :          :   :    |:   ::   :
Cp     770 ATCATGTCTAGCAACTCCTTCTCCCCAACTTGGTCAGGACATCATCATGGAGGCCTCG 711

Db     377 tqnvvmdknndrntdnvnwamqdndsqdnnnaahysqanknnwwtqrnnnwvkqannsdn 436
                 :: ::     :  |:  :          :   :    :          |:   :|    :     :       |          ::||    ::
Cp     710 GTGAAGGGCGCACTGGGGCGGCTGATCAGCACATGGGGCGGCTCAGCCTCCAGGAGGGTG 651

Db     437 nncandnddnsed-ktnrstnnnvanqtqntnnmnavssnnnerknnnknnasmnwrnr 495
                 ||  :   :     :|:  :        :|  |   :       ||                   :    :::        :     :     :        :           :
Cp     650 AGCACTAGCTGCTCGGCCTCAGGCGTCTAGCAGTAGCTCTCGTACCTGGCTCCGTGG 591

Db     496 wnnnnnqnsnryhkqngsrntnsnrgssyqsnmtahgkynnnantqhnkgnvvankhvnk 555
                              ::: |    :         :    :     :   |     :     :   ::                 |:          :         :::
Cp     590 CCGCCACTTCTCTTGGCCTTGCCGGCACAGTGCAGCTGCTCGTCGGCACTTCTCTGTCTC 531

Db     556 krnnntrnvnnnnkhmrdvnnnhntrnngacndnnnncnvtnycnrgsnngnnnndsnnn 615
                :        :   :            ::  ::          :  |             |  :            |   :     :      :      :       :
Cp     530 CGCACAAGGCGCTACCCACATCTCTCTCCGGCACCCACACTTCACCATTCCCACTTCC 471

Db     616 dwnnrysnnndnvkqmannhnsnnsshqsnkssnnvvdsrnvnkntdyqnasnrstannd 675
              :::  ::        :  ::   :        :   :        :::|:      ::  |      :::    :    :   |::          :    :   |
```

```
Cp     470 TAACACTTCCGAAGTC-GGCAGG-CCTGGCAGCTCTTGCGCCGGTTTTTATCGATTGTAC 413

Db     676 dnnanyakknntannnnsqnnnnttgmnaadvysnqnnnnnnanrsgnnynnqndnsnkn 735
              :       :  ::                  :          |          |  ::::                                    :          :       :
Cp     412 ACTGATTTGTAGCTGGACAAATATAATCATTATGTCCTTGAATGCTTCTTTTAAAAAAGG 353

Db     736 nvnkvrngnrnynrnsndrtnnnnnnvnnmnrcwandnanrndnqnnkqnnrrnnknqgt 795
                :  ::          :   :     ::              :         :      |  :            :  |     :             :        :       '
Cp     352 CCTTACATCCTTCATACCACCAGACTCCATAGTGATATCCCGATCCGTAATCGCTGCAGA 293

Db     796 snndnnnnrmnnyannnnknvnnrtnaynnnkrkanannynnnnhsvannnkrqntvnan 855
              :             ::        |              :          |                :::|          :             :::           :   |     :
Cp     292 CAGCCCAGAAGTGAGCATCCCTCTTTGAACCTGGACCAGTAAC-ACCCCTCCCGCAACGG 234

Db     856 andsvtnynsdnvqntansanstnmnvvtnnndnytcndannndnndvykvntnqdaymvv 915
                  :::   :    :    :                    :              :|       :   :|        |     :     ::::       |      :|  :::
Cp     233 -TTCCCACTAACCTTCCTTTTCAGTGTCTCTCTGTTTACAGGTAAGGTGTGTTCTAGCGA 175

Db     916 sqnnqrnanrhannnarnanannddavssnrnrhrnhdnnrnnqvhtqnnnaavvnnkm 974
                       |           . . |            .  |                 .          .         :        .   :   ::       :   :   |            :|  |::|   :
Cp     174 TCTTGCTTCACACCAGGGACTCTTTTGAGGTTCCGCATACAGATGTGATAACTGGCGATG 115

Db     975 nrynnnaltvntasrmnsngnankhhvssttkdandnnqnnnnnrqdvnmkahqk 1030
                 . . |          .    .    . |  . . . .                          :   : | |  : :         :           |                 | : :        : | : | :
Cp     114 GACCACTAAAGGAGAAAGGTGCCCAGGTGTTGGCCACAACACATTTGGGCTTGTGG 59

RESULT   12
ID   Q76269 standard; DNA; 6244 BP.
AC   Q76269;
DT   17-JUL-1995  (first entry)
DE   PhCMV*-1-controlled progesterone receptor
KW   tTA, transactivator, tetracycline-controllable transactivator;
KW   conditional inactivation, homologous recombination, gene expression;
KW   gene regulation; gene therapy; tetracycline-resistance; tetR;
KW   transgenic animal, PhCMV*-1, promoter, tetO, CMV;
KW   progesterone receptor; ds.
OS   Human cytomegalovirus Towne.
PN   WO9429442-A.
PD   22-DEC-1994.
PF   14-JUN-1994; U06734.
PR   14-JUN-1993; US-076327.
PA   (BADI ) BASF AG.
PI   Bujard H, Gossen M, Salfeld JG, Voss JW;
DR   WPI; 95-036472/05.
PT   Regulatory systems using tetracycline-controllable transactivator
PT   (tTA) - useful for conditional inactivation or modulation of
PT   gene expression in a host cell or animal
PS   Disclosure; Page 54-58; 103pp; English.
CC   The minimal promoter PhCMV*-1 is based on the human CMV
CC   intermediate-early promoter (with enhancer region removed), fused
CC   to a tetracycline operator, tetO.  The promoter is used to
CC   express tTA transactivator in host cells.  In clone pUHD BGR3
CC   (given in Q76269), cDNA for rabbit progesterone receptor is under
CC   the control of PhCMV*-1.
SQ   Sequence   6244 BP;    1483 A;    1749 C;    1610 G;    1402 T;

  Query Match             5.0%;  Score 63;  DB 13;  Length 6244;
  Best Local Similarity  66.3%;  Pred. No. 3.17e-25;
  Matches  128;  Conservative    0;  Mismatches   65;  Indels    0;  Gaps    0;

Db    2365 tctgtgggagatgaagatcaggctgtcattatggtgtcctcactctgtgggagtctcaaga 2424
              ||||   |  |||  |  |||||  ||   |   |||   ||||| |||          |  ||||       |  ||||||
Qy     293 TCTGCAGCGATTAGCATCGGATATCACTATGGAGTCTGGTCCTGTGAAGGATCTAAGG 352

Db    2425 tctttctttaaaagggcaatggaagggcagcataactatttatgtgctgggggggatgact 2484
                |||  ||||||||         ||     ||||  ||     ||  |   |||   |   |||  |  |     |   ||||  |  |
Qy     353 CCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCAGT 412

Db    2485 gcattgttgatgaaatccgcaggaaaaactgcccaggcatgtcgccttagggaatgctatc 2544
              |  |      |  |||||||  |||     |   ||  |  |||||  |||  ||  ||  |||  |  ||||||  |
Qy     413 GTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTACG 472
```

```
Db    2545 aagctggcatggt 2557
           ||| || |||||
Qy     473 AAGTGGGAATGGT 485


RESULT   13
ID   T06872 standard; DNA; 6244 BP.
AC   T06872;
DT   13-MAR-1996  (first entry)
DE   PhCMV*-1 promoter and rabbit progesterone receptor.
KW   Transactivator; tTA; tet operator; tetO; gene expression;
KW   tetracycline-responsive promoter; PhCMV*-1; human cytomegalovirus;
KW   transcription activation; eukaryotic cell; progesterone receptor;
KW   ds; ss.
OS   Chimeric synthetic.
OS   Chimeric Oryctolagus cuniculus
PN   US5464758-A.
PD   07-NOV-1995.
PF   14-JUN-1993; 076726.
PR   14-JUN-1993; US-076726.
PA   (BUJA/) BUJARD H
PA   (GOSS/) GOSSEN M.
PI   Bujard H,  Gossen M;
DR   WPI; 95-392612/50.
PT   Polynucleotide encoding trans-activator fusion protein contg. tet
PT   repressor  - used to control expression of gene regulated by minimal
PT   promoter linked to tet operon, and vectors and cells where gene
PT   expression is regulated by tetracycline
PS   Disclosure; Fig 9A-C; 37pp; English.
CC   cDNA coding for the rabbit progesterone receptor was placed under
CC   control of the PhCMV*-1 promoter (see T06869), resulting in
CC   tetracycline-regulated expression vector pUHDpgr-3 contg. the
CC   sequence given in T06872.  The construct was transfected into
CC   eukaryotic cells expressing the tTA transactivator (T06867) to
CC   allow tetracycline-regulated expression of the receptor.
SQ   Sequence   6244 BP;    1483 A;    1750 C;    1609 G;    1402 T;

  Query Match             5.0%;  Score 63;  DB 17;  Length 6244;
  Best Local Similarity  66.3%;  Pred. No. 3.17e-25;
  Matches  128;  Conservative    0;  Mismatches  65;  Indels    0;  Gaps    0;

Db    2365 tctgtggggatgaagcatcaggctgtcattatggtgtcctcacctgtgggagctgtaagg 2424
            |||  | ||  | ||||| || | ||| ||||| |||     | ||||    | |||||||
Qy     293 TCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATGTAAGG 352

Db    2425 tcttctttaaaggggcaatgagaggcagcataactatttatgtgactggaagaaatgact 2484
             ||| ||||||||      ||   |||| ||   || | ||| | |||  | |  | |||| | |
Qy     353 GCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCTAGCTACAAATCAGT 412

Db    2485 gcattgttgatcaaatccgcagggaaaactgcctggcgtgtcgccttagaaagtgctgtc 2544
            | |    | |||||||| |||    | || | ||||| ||| || || ||| | |||||| |
Qy     413 GTACAATCGATAAAAACGGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTACG 472

Db    2545 aagctggcatggt 2557
           ||| || |||||
Qy     473 AAGTGGGAATGGT 485


RESULT   14
ID   Q79934 standard; cDNA; 2992 BP.
AC   Q79934;
DT   01-AUG-1995  (first entry)
DE   Human liver hap cDNA.
KW   RAR-beta; retinoic acid receptor-beta; hap gene; hepatoma;
KW   hepatocellular carcinoma; hepatocyte; liver; hepatitis B virus;
KW   ss
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   misc_signal     119..121
FT   /*tag=  a
FT   /note= "untranslated start codon"
FT   misc_signal     126..128
FT   /*tag=  b
FT   /note= "untranslated start codon"
FT   misc_signal     178..180
FT   /*tag=  c
FT   /note= "untranslated start codon"
FT   misc_signal     237..239
FT   /*tag=  d
FT   /note= "untranslated start codon"
FT   CDS             322..1668
FT   /*tag=  e
FT   /product= RAR-beta
FT   polyA_signal    1954..2959
FT   /*tag=  f
PN   US5376530-A.
PD   27-DEC-1994.
PF   16-DEC-1987; 133687.
PR   16-DEC-1987; US-133687.
PR   17-DEC-1987; US-134130.
PR   20-JUN-1988; US-209009.
PR   30-NOV-1988; US-278136.
PR   30-MAR-1989; US-330405.
PR   21-AUG-1991; US-751612.
PR   30-MAR-1992; US-860577.
PR   11-DEC-1992; US-989902.
PR   22-JUL-1993; US-095706.
PA   (INSP ) INST PASTEUR.
PI   Brand N,  Chambon P,  De The HB,  Dejean A,  Krust A;
PI   Marchio A, Petkovich M,  Tiollais P;
DR   WPI; 95-043461/06.
DR   P-PSDB; R66764.
PT   Detection of transformed hepatocyte(s)  - by detecting a 2.5 kb
PT   mRNA species in hepatic tissue using DNA encoding a retinoic acid
PT   receptor-beta fragment.
PS   Disclosure; Fig.2A-2B; 35pp; English.
CC   A hepatitis B virus integration in a DNA fragment linked to
CC   hepatocellular carcinoma was used as a probe to clone cDNA from a
CC   human liver cDNA library.  The cellular gene, named hap, encoded a
CC   retinoic acid receptor identified as RAR-beta.
SQ   Sequence   2992 BP;    899 A;    652 C;    600 G;    841 T;

  Query Match             4.7%;  Score 59;  DB 13;  Length 2992;
  Best Local Similarity  64.8%;  Pred. No. 1.54e-22;
  Matches  129;  Conservative    0;  Mismatches  70;  Indels    0;  Gaps    0;

Db     561 ctgtttggtctgcccaggagaaatatcagggtaccactatggggtcagccctgtgaaggg 620
           ||||   ||||||   ||  |  |||| || || |||||||| ||| |   | ||||  ||
Qy     285 CTGCGTTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGG 344

Db     621 atgtaaggcttttttcagcagaagtattcagaagaatatgatttacacttgtcatccaga 680
           |||||||| ||||||     ||||| |||||       |||     ||| | |||||
Qy     345 ATGTAAGGCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCTAGCTAC 404

Db     681 taagaactgtgttattaataaagtcaccaggaatcgatgccaatactgtcgactccagaa 740
             ||  | |||    ||  |||||  |      | ||  | |||||    ||| |||||| | |||
Qy     405 AAATCAGTGTACAATCGATAAAAACGGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAA 464

Db     741 gtgctttgaagtgggaatg 759
           |||  |  ||||||||||||
Qy     465 GTGTTACGAAGTGGGAATG 483


RESULT   15
ID   Q65572 standard; cDNA; 2992 BP.
AC   Q65572;
DT   11-JAN-1995  (first entry)
DE   Human liver hap cDNA.
KW   Liver; hap; retinoic acid receptor; steroid; thyroid; hormone;
KW   hepatoma; retinoid; antibody; ss.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   CDS             322..1668
FT   /*tag=  a
FT   /product= RAR-beta retinoic acid receptor.
FT   misc_difference 916..918
FT   /*tag=  b
FT   /transl_except= GCT encodes Glycine.
FT   misc_difference 1372..1374
FT   /*tag=  c
FT   /transl_except= GGA encodes Alanine.
PN   US5317090-A.
PD   31-MAY-1994.
PF   16-DEC-1987; 133687.
PR   16-DEC-1987; US-133687.
PR   17-DEC-1987; US-134130.
PR   20-JUN-1988; US-209009.
PR   30-NOV-1988; US-278136.
PR   30-MAR-1989; US-330405.
PR   21-AUG-1991; US-751612.
PR   30-MAR-1992; US-860577.
PR   11-DEC-1992; US-989902.
PA   (INSP ) INST PASTEUR.
PI   Brand N,  Chambon P,  De THE HB,  Dejean A,  Krust A;
PI   Marchio A, Petkovich M,  Tiollais P;
DR   WPI; 94-176333/21.
PT   Antibody specific for retinoic acid receptor-beta - useful for
PT   detecting, quantifying and identifying agonists and antagonists
PT   of retinoid activity.
PS   Disclosure; Figure 2; 35pp; English.
CC   The hap cDNA encodes a retinoic acid receptor called RAR-beta.  The
CC   hap gene is transcribed at low level in most human tissues, but the
CC   gene is overexpressed in prosate and kidney.  Six out of seven
CC   hepatoma or hepatoma-derived cell lines express a small hap
CC   transcript which is undetectable in normal adult and foetal livers
CC   but present in all non-hepatic tissues tested.
SQ   Sequence   2992 BP;    899 A;    653 C;    599 G;    841 T;

  Query Match             4.7%;  Score 59;  DB 11;  Length 2992;
  Best Local Similarity  64.8%;  Pred. No. 1.54e-22;
  Matches  129;  Conservative    0;  Mismatches  70;  Indels    0;  Gaps    0;

Db      561 ctgcttcgtctgccaggacaaatcatcagggtaccactatggggtcagcgcctgtgaggg 620
            ||||   ||||||   ||  |  |||| || || |||||||| ||| |  | ||||| ||
Qy      285 CTGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGG 344

Db      621 atgtaagggctttttccgcagaagtattcagaagaatatgatttacacttgtcacccaga 680
            |||||||| ||||||    ||||| |||||     |||    ||| | |||||   |
Qy      345 ATGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTAC 404

Db      681 taagaactgtgttattaataaagtcaccaggaatcgatgccaatactgtcgactccagaa 740
             ||  | |||   ||  |||||  |    | ||  | |||||   ||| ||||| | |||
Qy      405 AAATCAGTGTACAATCGATAAAAACGGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAA 464

Db      741 gtgctttgaagtgggaatg 759
            ||| |  ||||||||||||
Qy      465 GTGTTACGAAGTGGGAATG 483

Search completed: Thu Dec 18 18:56:39 1997
Job time : 147 secs.
```

```
**********************************************************************

                      M P S R C H                            (TM)

**********************************************************************

           Release 2.1D John F. Collins, Biocomputing Research Unit
           Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
                 Distribution rights by IntelliGenetics, Inc.

MPsrch_nn   n.a. - n.a. database search, using Smith-Waterman algorithm

Run on:          Thu Dec 18 18:35:59 1997;  MasPar time 1060.11 Seconds
                                            1296.037 Million cell updates/sec
Tabular output not generated.

Title:           >US-08-826-361A-2
Description:     (1-1251) from US08826361A seq
Perfect Score:   1251
N.A. Sequence:      1 ATGAATTACAGCATTCCTAG            ATCCCAGCCATCCGAGCCA 1251
     Comp.            TACTTAATGTCGTAAGGATC            TAGGGTCGGTAGGCTCGGT

Scoring table:   TABLE default
                 Gap 6

Nmatch   STD :   Pbase 0;  Query 0

Searched:        362067 seqs  549138275 bases x 2

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database:        embl-new3
                   1:BCT 2:FUN 3:GEN1 4:GEN2 5:HTG1 6:HTG2 7:HUM 8:INV
                   9:ORG 10:MAM 11:VRT 12:PLN 13:PRO 14:ROD 15:SYN 16:UNC
                   17:VIR
Database:        genbank99
                   18:BCT1 19:BCT2 20:BCT3 21:BCT4 22:BCT5 23:BCT6 24:BCT7
                   25:BCT8 26:BCT9 27:BCT10 28:BCT11 29:GEN1 30:GEN2
                   31:GEN3 32:HTG1 33:HTG2 34:HTG3 35:INV1 36:INV2 37:INV3
                   38:INV4 39:INV5 40:INV6 41:INV7 42:INV8 43:INV9 44:INV10
                   45:INV11 46:MAM1 47:MAM2 48:MAM3 49:VRT1 50:VRT2 51:VRT3
                   52:VRT4 53:PAT1 54:PAT2 55:PAT3 56.PAT4 57.PAT5 58.PHG
                   59:PLN1 60:PLN2 61:PLN3 62:PLN4 63:PLN5 64.PLN6 65.PLN7
                   66:PLN8 67:PLN9 68:PLN10 69:PLN11 70:PRI1 71:PRI2
                   72:PRI3 73:PRI4 74:PRI5 75:PRI6 76:PRI7 77:PRI8 78:PRI9
                   79:PRI10 80:PRI11 81:PRI12 82:PRI13 83:PRI14 84:PRI15
                   85:ROD1 86:ROD2 87:ROD3 88:ROD4 89:ROD5 90:ROD6 91:ROD7
                   92:ROD8 93:STS 94:SYN 95:UNA 96:VRL1 97:VRL2 98:VRL3
                   99:VRL4 100:VRL5 101:VRL6 102:VRL7 103:VRL8 104:VRL9
                   105:VRL10
Database:        genbank-new3
                   106:BCT 107:GEN1 108:GEN2 109:HTG1 110:HTG2 111:INV
                   112:MAM 113:VRT 114:PHG 115:PLN 116:PRI1 117:PRI2
                   118:ROD 119:SYN 120:UNA 121:VRL
Database:        u-embl50_99
                   122:part1

Statistics:      Mean 11.299;  Variance 4.407;  scale 2.564

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                               SUMMARIES
```

```
Result          Query
   No.  Score  Match Length DB  ID              Description                Pred. No.
------------------------------------------------------------------------------------------
     1   1247   99.7   1560 79  HSRNAERB        H.sapiens mRNA for es   0.00e+00
     2    841   67.2   2555 52  RNU57439        Rattus norvegicus est   0.00e+00
     3    261   20.9    331 79  CJERB           C.jacchus mRNA for es   6.96e-216
     4    170   13.6   1729 51  OPZMER          Killifish mRNA for es   9.42e-140
     5    155   12.4   1985 51  GGERR           Chicken mRNA for estr   1.61e-124
     6    149   11.9   2792 52  ZFIER           Poephila guttata estr   5.75e-118
     7    148   11.8   1466 47  OAESTREC        O.aries mRNA for estr   6.19e-117
     8    140   11.2   1662 52  SSOESREC        S.salar mRNA for oest   1.06e-108
     9    140   11.2   1788 48  SSESTREC        S.scrofa mRNA for est   1.06e-108
    10    140   11.2   1880 91  RNESTROR        Rat mRNA for estrogen   1.06e-108
    11    140   11.2   1991 91  RNERR           Rat mRNA for estrogen   1.06e-108
    12    140   11.2   2097 92  S82129          estrogen receptor [ra   1.06e-108
    13    136   10.9   2106 79  HSU47678        Human 90 kDa estrogen   1.35e-104
    14    135   10.8   1278 52  SMOERA          Rainbow trout estroge   1.43e-103
    15    135   10.8   2092 55  I08538          Sequence 1 from Paten   1.43e-103
    16    135   10.8   2092 80  HUMERMCF        Human estrogen recept   1.43e-103
    17    135   10.8   4963 56  I15368          Sequence 16 from pate   1.43e-103
    18    135   10.8   6450 74  HSERR           Human mRNA for oestro   1.43e-103
    19    132   10.6    907 51  S79925          estrogen receptor-rep   1.67e-100
    20    132   10.6   2078 88  MUSESTRU        Mouse estrogen recept   1.67e-100
    21    133   10.6   2619 80  HUMERVBCA       Human estrogen recept   1.59e-101
    22    126   10.1   2449 52  XELESTRREC      Xenopus laevis estrog   2.19e-94
    23    126   10.1   2479 52  XFLESTRC        Xenopus laevis estrog   2.19e-94
    24    126   10.1   2533 52  XELERB          Xenopus laevis estrog   2.19e-94
    25    113    9.0    586 79  HSZ75126        H.sapiens mRNA for es   3.15e-81
    26    109    8.7    335 47  OAU30299        Ovis aries estrogen r   3.28e-77
    27     92    7.4    477 46  BTU64962        Bos taurus estrogen r   2.48e-60
    28     78    6.2    323 74  HSEST           H. sapiens mRNA for e   9.75e-47
    29     78    6.2    323 117  HSEST           H. sapiens mRNA for e  9.75e-47
    30     77    6.2    400 46  BTERER          B.taurus mRNA for est   5.85e-46
    31     74    5.9    365 51  OAESREXN4       O.aureus estrogen rec   6.42e-43
    32     71    5.7   3460 57  I25343          Sequence 1 from paten   4.45e-40
    33     68    5.4   2153 76  HSSTHOR2        Human mRNA for steroi   2.94e-37
    34     68    5.4   2153 55  I08706          Sequence 13 from Pate   2.94e-37
    35     66    5.3   1344 122  MMERR2GEN       M.musculus mRNA for e  2.16e-35
    36     66    5.3   1344 92  S82458          mERR-2=estrogen respo   2.16e-35
    37     64    5.1   2265 85  MMGRLA          Mouse mRNA for glucoc   1.55e-33
    38     62    5.0   2376 11  SOU87953        Saguinus oedipus gluc   1.08e-31
    39     62    5.0   2376 116  SOU87953        Saguinus oedipus gluc  1.08e-31
    40     62    5.0   2376 10  ANU87952        Aotus nancymaae gluco   1.08e-31
    41     62    5.0   2376 116  ANU87952        Aotus nancymaae gluco  1.08e-31
    42     62    5.0   2575 85  MMGCRR          Mouse mRNA for glucoc   1.08e-31
    43     63    5.0   2931 49  RABPGR          Rabbit progesterone r   1.30e-32
    44     63    5.0   6244 56  I15367          Sequence 15 from pate   1.30e-32
    45     61    4.9   1641 48  TBGCR           T.belangeri mRNA for    8.96e-31
```

```
                               ALIGNMENTS

RESULT     1
LOCUS       HSRNAERB     1560 bp    RNA              PRI       05-SEP-1996
DEFINITION  H.sapiens mRNA for estrogen receptor.
ACCESSION   X99101
NID         g1518262
KEYWORDS    estrogen receptor.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1560)
  AUTHORS   Mosselman,S., Polman,J. and Dijkema,R.
  TITLE     ER beta: identification and characterization of a novel human
            estrogen receptor
  JOURNAL   FEBS Lett. 392 (1), 49-53 (1996)
  MEDLINE   96354875
REFERENCE   2  (bases 1 to 1560)
  AUTHORS   Mosselman,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-JUL-1996) S. Mosselman, N.V.Organon, Biotechnology
            and Biochemistry, PO box 20 Oss, Molenstraat 110 Oss, 5340 BH,
            NETHERLANDS
FEATURES             Location/Qualifiers
     source          1..1560
                     /organism="Homo sapiens"
                     /tissue_type="testis"
                     /dev_stage="adult"
                     /clone_lib="Clontech HL1010B"
                     /clone_lib="Clontech; 7414-1"
     CDS             19..1452
                     /codon_start=1
                     /product="estrogen receptor beta"
                     /db_xref="PID:e255390"
                     /db_xref="PID:g1518263"
                     /translation="MNYSIPSNVTNLEGGPGRQTTSPNVLWPTPGHLSPLVVHRQLSH
                     LYAEPQKSPWCEARSLEHTLPVNRETLKRKVSGNRCASPVTGPGSKRDAHFCAVCSDY
                     ASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRKCYEVG
                     MVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQLVLT
                     LLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFDQVR
                     LLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSRF
                     RELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKSG
                     ISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNAHVLRGC
                     KSSITGSECSPAEDSKSKEGSQNPQSQ"
BASE COUNT      367 a    408 c    445 g    340 t
ORIGIN

  Query Match              99.7%;  Score 1247;  DB 76;  Length 1560;
  Best Local Similarity   100.0%;  Pred. No. 0.00e+00;
  Matches  1247;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Db         19 atgaattacagcattcccagcaatgtcactaacttggaaggtgggcctggtcggcagacc 78
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy          1 ATGAATTACAGCATTCCCAGCAATGTCACTAACTTGGAAGGTGGGCCTGGTCGGCAGACC 60

Db         79 acaagcccaaatgtgttgtggccaacacctgggcacctttctcctttagtggtccatcgc 138
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy         61 ACAAGCCCAAATGTGTTGTGGCCAACACCTGGGCACCTTTCTCCTTTAGTGGTCCATCGC 120

Db        139 cagttatcacatctgtatgcggaacctcaaaagagtccctggtgtgaagcaagatcgcta 198
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        121 CAGTTATCACATCTGTATGCGGAACCTCAAAAGAGTCCCTGGTGTGAAGCAAGATCGCTA 180

Db        199 gaacacaccttacctgtaaacagagagacactgaaaaggaaggttagtgggaaccgttgc 258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        181 GAACACACCTTACCTGTAAACAGAGAGACACTGAAAAGGAAGGTTAGTGGGAACCGTTGC 240

Db        259 gccagccctgttactggtccaggttcaaagagggatgctcacttctgcgctgtctgcagc 318
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        241 GCCAGCCCTGTTACTGGTCCAGGTTCAAAGAGGGATGCTCACTTCTGCGCTGTCTGCAGC 300

Db        319 gattacgcatcgggatatcactatggagtctggtcgtgtgaaggatgtaaggcctttttt 378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        301 GATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATGTAAGGCCTTTTTT 360

Db        379 aaaagaagcattcaaggacataatgattatatttgtccagctacaaatcagtgtacaatc 438
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        361 AAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCAGTGTACAATC 420

Db        439 gataaaaaccggcgcaagagctgccaggcctgccgacttcggaagtgttacgaagtggga 498
              ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        421 GATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTACGAAGTGGGA 480

Db        499 atggtgaagtgtggctcccggagagagagatgtgggtaccgccttgtgcggagacagaga 558
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        481 ATGGTGAAGTGTGGCTCCCGGAGAGAGAGATGTGGGTACCGCCTTGTGCGGAGACAGAGA 540

Db        559 agtgccgacgagcagctgcactgtgccggcaaggccaagagaagtggcggccacgcgccc 618
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        541 AGTGCCGACGAGCAGCTGCACTGTGCCGGCAAGGCCAAGAGAAGTGGCGGCCACGCGCCC 600

Db        619 cgagtgcgggagctgctgctggacgccctgagccccgagcagctagtgctcaccctcctg 678
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        601 CGAGTGCGGGAGCTGCTGCTGGACGCCCTGAGCCCCGAGCAGCTAGTGCTCACCCTCCTG 660

Db        679 gaggctgagccgccccatgtgctgatcagccgccccagtgcgcccttcaccgaggcctcc 738
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        661 GAGGCTGAGCCGCCCCATGTGCTGATCAGCCGCCCCAGTGCGCCCTTCACCGAGGCCTCC 720

Db        739 atgatgatgtccctgaccaagttggccgacaaggagttggtacacatgatcagctgggcc 798
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        721 ATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTACACATGATCAGCTGGGCC 780

Db        799 aagaagattcccggctttgtggagctcagcctgttcgaccaagtgcggctcttggagagc 858
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        781 AAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAAGTGCGGCTCTTGGAGAGC 840

Db        859 tgttggatggaggtgttaatgatggggctgatgtggcgctcaattgaccaccccggcaag 918
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        841 TGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCAATTGACCACCCCGGCAAG 900

Db        919 ctcatctttgctccagatcttgttctggacagggatgaggggaaatgcgtagaaggaatt 978
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        901 CTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGGAAATGCGTAGAAGGAATT 960

Db        979 ctggaaatctttgacatgctcctggcaactacttcaaggtttcgagagttaaaactccaa 1038
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        961 CTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTTCGAGAGTTAAAACTCCAA 1020

Db       1039 cacaaagaatatctctgtgtcaaggccatgatcctgctcaattccagtatgtaccctctg 1098
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1021 CACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAATTCCAGTATGTACCCTCTG 1080

Db       1099 gtcacagcgacccaggatgctgacagcagccggaagctggctcacttgctgaacgccgtg 1158
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1081 GTCACAGCGACCCAGGATGCTGACAGCAGCCGGAAGCTGGCTCACTTGCTGAACGCCGTG 1140

Db       1159 accgatgctttggtttgggtgattgccaagagcggcatctcctcccagcagcaatccatg 1218
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1141 ACCGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATCTCCTCCCAGCAGCAATCCATG 1200

Db       1219 cgcctggctaacctcctgatgctcctgtcccacgtcaggcatgcgag 1265
              |||||||||||||||||||||||||||||||||||||||||||||||
Qy       1201 CGCCTGGCTAACCTCCTGATGCTCCTGTCCCACGTCAGGCATGCGAG 1247


RESULT      2
LOCUS       RNU57439     2555 bp    mRNA            ROD       12-JUN-1996
DEFINITION  Rattus norvegicus estrogen receptor beta mRNA, complete cds.
ACCESSION   U57439
NID         g1373280
KEYWORDS    .
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
            Murinae; Rattus.
REFERENCE   1  (bases 8 to 2555)
  AUTHORS   Kuiper,G.G., Enmark,E., Pelto-Huikko,M., Nilsson,S. and
            Gustafsson,J.A.
  TITLE     Cloning of a novel receptor expressed in rat prostate and ovary
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 93 (12), 5925-5930 (1996)
  MEDLINE   96234066
REFERENCE   2  (bases 1 to 2555)
  AUTHORS   Enmark,E.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-MAY-1996) Eva L. Enmark, Center for Biotechnology,
            NOVUM, Karolinska, Institute, NOVUM, Huddinge, S-141 57, Sweden
FEATURES             Location/Qualifiers
     source          1..2555
                     /organism="Rattus norvegicus"
                     /strain="Sprague Dawley"
                     /sex="male"
                     /tissue_type="prostate"
     CDS             417..1874
                     /note="nuclear receptor superfamily member"
                     /codon_start=1
                     /product="estrogen receptor beta"
                     /db_xref="PID:g1373281"
                     /translation="MTFYSPAVMNYSVPGSISNLDGGPVRLSISPNVLWPTSGHLSPL
                     ATHCQSSLLYAEPQKSPWCEARSLEHTLPVNRETLKRKLSGSSCASPVTSPNAKRDAH
                     FCPVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACR
                     LRKCYEVGMVKCGSRRERCGYRIVRRQRSSSEQVHCLSKAKRNGGHAPRVKELLLSTL
                     SPEQLVLTLLEAEPPNVLVSRPSMPFTEASMMMSLTKLADKELVHMIGWAKKIPGFVE
                     LSLLDQVRLLESCWMEVLMVGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDM
                     LLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLASANQEAESSRKLTHLLNAVTDAL
                     VWVIAKSGISSQQQSVRLANLLMLLSHVRHISNKGMEHLLSMKCKNVVPVYDLLLEML
                     NAHTLRGYKSSISGSECSSTEDSKNKESSQNLQSQ"
     misc_feature    726..922
                     /note="encodes DNA binding domain"
BASE COUNT      630 a    652 c    660 g    613 t
ORIGIN

  Query Match             68.8%;  Score 861;  DB 92;  Length 2555;
  Best Local Similarity   84.7%;  Pred. No. 0.00e+00;
  Matches 1051;  Conservative    0;  Mismatches  190;  Indels    0;  Gaps    0;

Db    441 atgaactacagtgttcccggcagcaccagtaacctggacggtgggcctgtccgactgagc 500
          ||||| ||||  ||||| |||    || |||| |||| ||||||||||  || | || |
Qy      1 ATGAATTACAGCATTCCCAGCAATGTCACTAACTTGGAAGGTGGGCCTGGTCGGCAGACC 60

Db    501 acaagcccaaatgtgctatggccaacttctgggcacctgtctcctttaacgacccattgc 560
          ||||||||||||||| | ||||||||   ||||||||| |||||||||  | ||||| ||
Qy     61 ACAAGCCCAAATGTGTTGTGGCCAACACTTGGGCACCTTTCTCCTTTAGTGCCCCATCGC 120

Db    561 caattatggctcctctatgcagaacctcaaaagagtccttgatgtgaagcaagatcacta 620
          || ||||  |  |  ||||| ||||||||||||||||| || |||||||||||||| |||
Qy    121 CAGTTATCACATCTGTATGCGGAACCTCAAAAGAGTCCCTGGTGTGAAGCAAGATCGCTA 180

Db    621 gagcacaccttacctgtaaacagagagacactgaaaaggaagcttagtgggagcagttgt 680
          || ||||||||||||||||||||| ||||||||||||||||| |||||||||| | |||| 
Qy    181 GAACACACCTTACCTGTAAACAGACAGACACTGAAAAGGAAGGTTAGTGGGAACCGTTGG 240

Db    681 gccagccctgttactagtccaaacgcaaagagggatgctcacttctgccccgtctgcagc 740
          ||||||||||||||| |||||     |||||||||||||||||| ||||||||||||||| 
Qy    241 GCCAGCCCTGTTACTGGTCCAGGTCAAAAGAGGGATGCTCACTTTTGCCCCGTCTGCAGT 300

Db    741 gattatgcatctgggtatcattacggcgtttggtcatgtgaaggatgtaaggcctttttt 800
          ||||| ||||| || ||||| || || || ||||| ||||||||||||||||||||||||
Qy    301 GATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATGTAAGGCCTTTTTT 360

Db    801 aaaagaagcattcaaggacataatgattatatctgtccagccacgaatcagtgtaccata 860
          |||||||||||||||||||||||||||||||| ||||||||| || ||||||||||| || 
Qy    361 AAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCAGTGTACAATC 420

Db    861 gacaagaaccggcgtaaaagctgccaggcctgccgacttcgcaagtgttatgaagtagga 920
          || || |||||||| || |||||||||||||||||||||| ||||||||| ||||| |||
Qy    421 GATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTACGAAGTGGGA 480

Db    921 atggtcaagtgtggatccaggagagaacggtgtgggtaccgtatagtgcggaggcagaga 980
          ||||| ||||||||| ||| ||||||||  | |||||||||||  | |||||||| ||||||
Qy    481 ATGGTGAAGTGTGGCTCCCGGAGAGAGAGATGTGGGTACCGCCTTGTGCGGAGACAGAGA 540

Db    981 agttctagcgagcaggtacactgcctgagcaaagccaagagaaacggtgggcatgcaccc 1040
          ||| |   ||||||| | ||||||     |||| ||||||||| ||  ||||| || || |||
Qy    541 AGTGCCGACGAGCAGCTGCACTGTGCCGGCAAGGCCAAGACAAGTGGTGGCCACGCGCCC 600

Db   1041 cgggtgaaggagctactgctgagcaccttgagtccagagcaactggtgctcaccctcctg 1100
          || |||  |||||| |||||||  | || |||| || ||||| || ||||||||||||||||
Qy    601 CGAGTGCGGGAGCTGCTGCTGGACGCCCTGAGCCCCGAGCAGCTAGTGCTCACCCTCCTG 660

Db   1101 gaggctgaaccacccaatgtgctggtgagccgtcccagcatgcccttcaccgaggcctcc 1160
          |||||||| || ||| |||||||| | ||||| |||||   |||||||||||||||||||
Qy    661 GAGGCTGAGCCGCCCCATGTGCTGATCAGCCGCCCCAGTGCGCCCTTCACCGAGGCCTCC 720

Db   1161 atgatgatgtccctcactaagctggcggacaaggaactggtgcacatgattggctgggcc 1220
          |||||||||||||| || ||| |||| |||||||||  |||| ||||||||  ||||||||
Qy    721 ATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTACACATGATCAGCTGGGCC 780

Db   1221 aagaaaatccctggctttgtggagctcagcctgttggaccaagtccggctcttagaaagc 1280
          ||||| || || |||||||||||||||||||||||| ||||||||| |||||||| || |||
Qy    781 AAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAAGTGCGGCTCTTGGAGAGC 840

Db   1281 tgctggatggaggtgctaatggtgggactgatgtggcgctccatcgaccaccccggcaag 1340
          ||||||||||||||| |||| | |||| ||||||||||||| || ||||||||||||||||
Qy    841 TGCTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCAATTGACCACCCCGGCAAG 900

Db   1341 ctcatttttgctccagatcttgttctggacagggatgaggggaaatgcgtagaaggaatt 1400
          ||||| ||||||||||||||  |||||||||||||| |||||||||||||||||||||||
Qy    901 CTCATCTTTGCTCCAGATCTCATTCTGGACAGGGATCAGGGGAAATGCGTAGAAGGAATT 960

Db   1401 ctggaaatctttgacatgctcctggcgacgacgtcaaggttccgtgagttaaaactccaa 1460
          |||||||||||||||||||||||||| || || |||||||| || |||||||||||||||
Qy    961 CTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTTCGAGAGTTAAAACTCCAA 1020

Db   1461 cacaaggagtatctctgtgtgaaggccatgatcctcctcaactccagtatgtaccccttg 1520
          ||||| || ||||| ||||| ||||||||||||||| |||| | ||||||||||||  ||
Qy   1021 CACAAAGAATATCTTTGTGTTAAGGCCATGATCCTGCTCAATTCCAGTATGTACCCTCTG 1080

Db   1521 gcttctgcaaaccaggaggcagaaagtagccgaaagctgacacacctactgaacgcggtg 1580
          |   | |  | |||||| || || || ||||| |||||||| | ||| | |||||| |||
Qy   1081 GTCACAGGGACCCAGGATGCTGACAGCAGCCGGAAGCTGGCTCACTTGCTGAACGCCGTG 1140

Db   1581 acagatgccctggtctgggtgattgcgaagagtggtatctcctcccagcagcagtcagtc 1640
          || |||||  |||| |||||||||||| ||||| || |||||||||||||||||| ||  | 
Qy   1141 ACCGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATCTCCTCCCAGCAGCAATCCATG 1200

Db   1641 cgactggccaacctcctgatgcttctttctcacgtcaggca 1681
          || ||||| |||||||||||||| || || |||||||||||
Qy   1201 CGCCTGGCTAACCTCCTGATGCTCCTGTCCCACGTCAGGCA 1241


RESULT      3
LOCUS       CJERB        332 bp    RNA             PRI       14-NOV-1996
DEFINITION  C.jacchus mRNA for estrogen receptor beta.
ACCESSION   Y09372
NID         g1669661
KEYWORDS    DNA binding; estrogen receptor; steroid hormone receptor;
            transactivation.
SOURCE      common marmoset.
  ORGANISM  Callithrix jacchus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Platyrrhini; Callitrichidae;
            Callithrix.
REFERENCE   1  (bases 1 to 332)
  AUTHORS   Gaughan,J.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 332)
  AUTHORS   Gaughan,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-NOV-1996) J. Gaughan, Medical Research Council,
            Centre for Reproductive Biology, 37 Chalmers Street, Edinburgh, EH3
            9EW, UK
COMMENT     Related entries: X99101 & U57439
            Related entries: X99101 & U57439.
FEATURES             Location/Qualifiers
     source          1..332
                     /organism="Callithrix jacchus"
                     /tissue_type="testis"
                     /tissue_type="ovary"
                     /tissue_type="prostate"
                     /clone="pGEM-T vector"
     CDS             18..>332
                     /function="steroid hormone receptor"
                     /codon_start=1
                     /product="estrogen receptor beta"
                     /db_xref="PID:e281463"
                     /db_xref="PID:g1669662"
                     /translation="MNYSIPSSVINLEAGPGRQTTSPNMLWSTPGHLSPLAVHHQLSH
                     LYAEPQKTPWCEARSLEHTLPVSRETLKRKVSGNHCASPVTGPSSKRDAHFCAVCSDY
                     AAE"
BASE COUNT       91 a     84 c     77 g     80 t
ORIGIN
```

```
  Query Match             20.9%;  Score 261;  DB 70;  Length 332;
  Best Local Similarity   92.2%;  Pred. No. 6.96e-236;
  Matches  285;  Conservative    0;  Mismatches   24;  Indels    0;  Gaps    0;

Db      18 atgaattacagcattcccagcagtgttactaacttggaagcggggcctggtcggcagatc 77
           |||||||||||||||||||||| ||| |||||||||||||   |||||||||||||||| |
Qy       1 ATGAATTACAGCATTCCCAGCAATGTCACTAACTTGGAAGGTGGGCCTGGTCGGCAGACC 60

Db      78 acaagcccaaatatgctgtggtcaacacctggccacctttctcctttagctgtccatcac 137
           |||||||||||| || ||||| |||||||||| ||||||||||||||||  ||||||| |
Qy      61 ACAAGCCCAAATGTGTTGTGGCCAACACCTGGGCACCTTTCTCCTTTAGTGGTCCATCGC 120

Db     138 cagttatcacatctgtatgcagaacctcaaaagactccctggtgtgaagcaagatcacta 197
           |||||||||||||||||||| |||||||||||||| ||||||||||||||||||||| |||
Qy     121 CAGTTATCACATCTGTATGCGGAACCTCAAAAGAGTCCCTGGTGTGAAGCAAGATCGCTA 180

Db     198 gaacataccttacctgtaagcagagagacactgaaaaggaaagttagtgggaaccattgt 257
           ||||| ||||||||||||| |||||||||||||||||||||| ||||||||||||| |||
Qy     181 GAACACACCTTACCTGTAAACAGAGAGACACTGAAAAGGAAGGTTAGTGGGAACCGTTGC 240

Db     258 gccagccctgttactggtccaggttcaaagagggatgctcacttctgtgcagtctgcagc 317
           |||||||||||||||||||| ||||||||||||||||||||||||||||| || ||||||||
Qy     241 GCCAGCCCTGTTACTGGTCCAGGTTCAAAGAGGGATGCTCACTTCTGGCGTGTCTGCAGC 300

Db     318 gattatgca 326
           ||||| |||
Qy     301 GATTACGCA 309


RESULT      4
LOCUS       ORZMER        1728 bp    mRNA             VRT       22-DEC-1994
DEFINITION  Killifish mRNA for estrogen receptor, complete cds.
ACCESSION   D28954
NID         g473547
KEYWORDS    estrogen receptor.
SOURCE      Oryzias sp. (strain d-rR) (library lambda ZAP II) liver cDNA to
            mRNA.
  ORGANISM  Oryzias sp.
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Actinopterygii; Neopterygii; Teleostei; Euteleostei;
            Acanthopterygii; Atherinomorpha; Cyprinodontiformes;
            Adrianichthyoidei; Oryziidae; Oryzias
REFERENCE   1  (bases 1 to 1728)
  AUTHORS   Okada,H.
  JOURNAL   Unpublished (1994)
COMMENT     Submitted (15-Mar-1994) to DDBJ by:
            Hisashi Okada
            Center for Gene Science
            Hiroshima University
            1-4-2 Kagamiyama, Higashi Hiroshima
            Hiroshima 724
            Japan
            Phone: 0824-24-6283
            Fax:   0824-22-7184.
FEATURES             Location/Qualifiers
     source          1..1728
                     /organism="Oryzias sp."
                     /strain="d-rR"
                     /tissue_type="liver"
                     /clone_lib="lambda ZAP II"
     CDS             1..1728
                     /gene="MER"
                     /codon_start=1
                     /product="estrogen receptor"
                     /db_xref="PID:g602326"
                     /translation="MYPEESRGSGGVLLWIFWKGRTIMPAPNPATIPLYSQSSIGYYS
                     APLETNGPPSEGSIQSLGSGPTSPIVFVPSSPRLSPFMHPPSHHYIETTSTPVYRSSH
                     QGASREDQCGSREDTCSLGELGAGAGAGGFEMAKDTRFCAVCSDYASGYHYGVWSCEG
                     CKAFFKRSIQGHNDYMCPATNQCTIDRNRRKSCQACRLRKCYEVGMMKGGVRKDRIRI
                     LRRDKRRIGVGDGDKVVKGQEHKIVHYDGRKRSSIGGGGGGGGRLSVTSIPPEQVLL
                     LLQGAEPPILCSRQKLSRPYTEVTMMTLLTSMADKELVHMIAWAKKLPGFLQLSLHDQ
                     VLLLESSWLEVLMIGLIWRSIHCPGKLIFAQDLILDRNEGDCVEGMTEIFDMLLATAS
                     RFRVLKLKPEEFVCLKAIILLNSGAFSFCTGTMEPLHNSAAVQSMLDTITDALIHYIS
                     QSGYLAQEQARPQAQLLLLLSHIRHMSNKGMEHLYSMKCKNKVPLYDLLLEMLDAHRL
                     HHPVRAPQSLSQVDRDPPSTSSGGGGIAPGSISASRGRIESPSRGPFAPSVLQYGGSR
                     PDCTPALQD"
BASE COUNT      379 a    513 c    496 g    340 t
ORIGIN

  Query Match             13.6%;  Score 170;  DB 51;  Length 1728;
  Best Local Similarity   73.2%;  Pred. No. 9.42e-140;
  Matches  268;  Conservative    0;  Mismatches   98;  Indels    0;  Gaps    0;

Db     883 ccgtacaccgaggtcaccatgatgacccctgctcaccagcatggcagacaaggagctggtc 942
           || | |||||||| | |||||||||       || ||||    |||| |||||||||| ||||
Qy     703 CCCTTCACCGAGGCCTCCATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTA 762

Db     943 cacatgatcgcctgggccaagaagctcccaggtttcctgcagctgtccctgcacgatcag 1002
           |||||||||   ||||||||||||| | || || ||| || ||||    ||||  |||  |
Qy     763 CACATGATCAGCTGGGCCAAGAAGATTCCGGGCTTTGTGGAGCTCAGCCTGTTCGACCAA 822

Db    1003 gtgctgctgctggagagctcgtggctggaggtgctcatgatcggcctcatttggaggtcc 1062
           |||| |||  |||||||||  ||| ||||||||| |||||| || || || ||| | ||
Qy     823 GTGCGGCTCTTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCA 882

Db    1063 atccactgtcccgggaagctcatctttgcacaagacctcatcctggacaggaatgagggg 1122
           ||  ||   ||||| |||||||||||||||| | ||| ||   | |||||||||| |||||
Qy     883 ATTGACCACCCCGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGG 942

Db    1123 gactgcgtggaaggcatgacggagatcttcgacatgctgctggccactgcttcccgcttc 1182
             | ||||| ||||| ||    ||| ||||| |||||||| ||||| ||| ||||  | ||
Qy     943 AAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTT 1002

Db    1183 cgtgtgctcaaactcaaacctgaggaattcgtctgcctcaaagctattattttactcaac 1242
           || | | | |||||| |||    | ||||    ||||   |||| || || ||   | |||| |
Qy    1003 CGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAAT 1062

Db    1243 tccggt 1248
           ||| ||
Qy    1063 TCCAGT 1068


RESULT      5
LOCUS       GGERR         2085 bp    RNA              VRT       12-SEP-1993
DEFINITION  Chicken mRNA for estrogen receptor.
ACCESSION   X03805
NID         g63378
KEYWORDS    DNA binding protein; estrogen receptor; receptor; steroid hormone
            receptor; unidentified reading frame.
SOURCE      chicken.
  ORGANISM  Gallus gallus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Archosauria; Aves; Neognathae; Galliformes;
            Phasianidae; Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 2085)
  AUTHORS   Krust,A., Green,S., Argos,P., Kumar,V., Walter,P., Bornert,J.M. and
            Chambon,P.
  TITLE     The chicken oestrogen receptor sequence: homology with v-erbA and
            the human oestrogen and glucocorticoid receptors
  JOURNAL   EMBO J. 5 (5), 891-897 (1986)
  MEDLINE   86247578
COMMENT     Data kindly reviewed (23-JUL-1986) by S. Green.
FEATURES             Location/Qualifiers
     source          1..2085
                     /organism="Gallus gallus"
     mRNA            1..2085
                     /note="ER mRNA"
     CDS             69..131
                     /note="open reading frame 1 (aa 1-20)"
                     /codon_start=1
                     /db_xref="PID:g63379"
                     /translation="MYCCFNMISAEPQNRFWCFF"
     CDS             177..200
                     /note="open reading frame 2 (aa 1-7)"
                     /codon_start=1
                     /db_xref="PID:g584490"
                     /translation="MFLAHCH"
     CDS             203..1972
                     /note="chicken estrogen receptor (aa 1-589)"
                     /codon_start=1
                     /db_xref="PID:g63380"
                     /db_xref="SWISS-PROT:P06212"
                     /translation="MTMTLHTKASGVTLLHQIQGTFLETLSRPQLKIPLERSLSDMYV
                     ESNKTGVFNYPEGATYDFGTTAPVYGSTTLSYAPTSESFGSSLAGFHSLNNVPTSPV
                     VFLQTAPQLSPFIHHHSQQVPYYLFNEQGSFGMREAAPPAFYRPSSDNPRHSIPEPMS
                     STNEKGSLSMESTKETRYCAVCNDYASGYHYGVWSCEGCKAFFKRSIQGHNDYMCPAT
                     NQCTIDKNRRKSCQACRLRKCYEVGMMKGGIRKDRRGGEMMKQKRQREEQDSPNGFAS
                     STELRAPTLWTSPLVVKHNKKNSPALSLTAEQMVSALLEAEPPIVYSEYDPNRPFNEA
                     SMMTLLTNLADRELVHMINWAKRVPGFVDLTLHDQVHLLECAWLEILMIGLVWRSMEH
                     PGKLIFAPNLLLDRNQGKCVEGMVEIFDMLLATAARFRMMNLQGEEFVCLKSIILLNS
                     GVYTFLSSTLKSLEERDYIHRVLDKITDTLIHLMAKSGLSLQQQHRRLAQLLLILSHI
                     RHMSNKGMEHLYNMKCKNVVPLYDLLLEMLDAHRLHAPAARSAAPMEEENRNQLTTAP
                     ASSHSLQSFYINSKEEESMQNTI"
     misc_feature    203..313
                     /note="aa 1-37 => highly conserved region A"
     misc_feature    314..721
                     /note="conserved region B (aa 38-173)"
     misc_feature    722..970
                     /note="highly conserved region C (aa 174-256) pot.
                     DNA-binding domain"
     misc_feature    971..1087
                     /note="conserved region D (aa 257-295)"
     misc_feature    1088..1840
                     /note="highly conserved region E (aa 296-546) certain
                     estrogen binding domain"
     misc_feature    1841..1969
                     /note="conserved region F (aa 547-589)"
BASE COUNT      582 a    519 c    480 g    504 t
ORIGIN

  Query Match             12.4%;  Score 155;  DB 50;  Length 2085;
  Best Local Similarity   65.2%;  Pred. No. 3.61e-124;
  Matches  364;  Conservative    0;  Mismatches  191;  Indels    3;  Gaps    2;

Db   1180 ccccaatagaccattcaacgaagcatctatgatgatctgttgaccaacttgcagacag 1239
            ||||| |    || |||| ||| || || |||||||      |||||||  | || ||||
Qy    693 CCCCAGTTGTGCCTTTAACGGAGGCCTTCATGATGATGTTTTTGATTAAGTTGGCCGACAA 752

Db   1240 aggattagtgcacatgatcaactgggcaaagagagttccaggatttgtggatttaacact 1299
             || || || |||||||||| |||||| ||||    |||| || ||||||||  | |  ||
Qy    753 GGAGTTGGTACACATGATCAACTGGGCCAAGAAGATTCCCGGCTTTGTGGAGCTCAGCCT 812

Db   1300 ccatgatcaggtccatctgctggaatgtgcctggttagagatattgatgatcggcttagt 1359
                || || || |  ||  ||||  |    ||| | ||| | || ||||| ||  |  |
Qy    813 GTTCGACCAAGTGCGGCTCTTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGAT 872

Db   1360 ctggcgtccatggaacaccaggaaagctttatttgaccctaatctattactggacag 1419
            ||||||| || || ||||| || |||||   | ||||| ||  ||||   | ||||||||
Qy    873 GTGGCGCTCAATTGACCACCCGGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAG 932

Db   1420 gaatcaagggaaatgtgtagagggcatggtggaaatctttgacatgctactggctactgc 1479
            | || | |||||||| ||||| || ||  ||||||||||||||||||||| ||||| |||
Qy    933 GGATGAGGGGAAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTAC 992

Db   1480 tgctcggtttcggatgatgaaccttcaaggggaggaatttgtgtgccttaagtccatcat 1539
            | |  |||||||    | | || || |||     | |||| | | ||  | ||| |||| ||
Qy    993 TTCAAGGTTTCGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGAT 1052

Db   1540 cctgctcaattctggtgtgtacactttttctttctagcaccttgaaatctctggaagagag 1599
            |||||||||||   || ||||| || |  |   |     |||  | |  ||    |  ||
Qy   1053 CCTGCTCAATTCCAGTATGTACCCTCTGGTCACAGCGACCCAGGATGCTGAC-AGCAGCC 1111

Db   1600 ggactatatccacggtgttctggacaaaatcacagatactctgatacacctaatggcaaa 1659
            |||    |   |    | ||  ||| ||     | || ||| || || |       | || || ||
Qy   1112 GGAAGCTGGCTCACTTG---CTGAACGCCGTGACCGATGCTTTGGTTTGGGTGATTGCCAA 1169

Db   1660 gtcaggtctttctctgcagcagcaacaccggcgactagctcagctcctccttatcctctc 1719
            |    ||  | ||     |||||||| |  ||| || ||| | |||||  |  |||| ||
Qy   1170 GAGCGGCATCTCCTCCCAGCAGCAATCCATGCGCCTGGCTAACCTCCTGATGCTCCTGTC 1229
```

```
Db   1720 tcacatcaggcatatgag 1737
            ||| |||||||  |||
Qy   1230 CCACGTCAGGCATGCGAG 1247


RESULT    6
LOCUS       ZFIER        2792 bp    mRNA            VRT       23-JUL-1996
DEFINITION  Poephila guttata estrogen receptor mRNA, complete cds.
ACCESSION   L79911
NID         g1449145
KEYWORDS    estrogen receptor.
SOURCE      Poephila guttata female adult hypothalamus-preoptic area cDNA to
            mRNA.
  ORGANISM  Poephila guttata
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Archosauria; Aves; Neognathae; Passeriformes;
            Estrildidae; Estrildinae; Poephila.
REFERENCE   1  (bases 1 to 2792)
  AUTHORS   Jacobs,E.C., Arnold,A.P. and Campagnoni,A.T.
  TITLE     Zebra finch estrogen receptor cDNA: cloning and mRNA expression
  JOURNAL   J. Steroid Biochem. Mol. Biol. 59(2) (1996) In press
FEATURES             Location/Qualifiers
     source          1..2792
                     /organism="Poephila guttata"
                     /dev_stage="adult"
                     /sex="female"
                     /tissue_type="hypothalamus-preoptic area"
     5'UTR           <1..132
     exon            <1..560
                     /note="Exon boundaries based on alignment to chicken ER"
                     /number=1
     misc_feature    127..237
                     /note="domain A"
     CDS             133..1896
                     /codon_start=1
                     /product="estrogen receptor"
                     /db_xref="PID:g1449146"
                     /translation="MTLHTKTSGVILLHQIQGTELETLSRPQLKIPLERSLSDMYVET
                     NKTGVFNYPEGATYDFGTTAPVYSSTTLSYAPTSESFGSSSLAGFHSLNSVPPSPVVF
                     LQIAPHWSPFIHHHSQQVPYYLENDQGSFGMREAAPPAFYRPNSDNRRHSIRERMSSA
                     NEKGSLSMESTKETRYCAVCNDYASGYHYGVWSCEGCKAFFKRSIQGHNDYMCPATNQ
                     CTIDKNRRKSCQACRLRKCYEVGMMKGGIRKDRRGGRVMKHKRQREEQDSRNGEASST
                     ELRAPTLWASPLVVKHNKKNSPALSLTAEQMVSALLEAEPPLVYSEYDPNRPFNEASM
                     MTLLTNLADRELVHMINWAKRVPGFVDLTLHDQVHLLECAWLEILMIGLVWRSMEHPG
                     KLLFAPNLIIDRNQGKCVEGMVEIFDMLLATAARFRMMNLQGEEFVCLKSIILLNSGV
                     YTFLSSTLKSLEEKDYIHRVLDKITDTLIHLMAKSGLSLQQQHRRLAQLLLILSHIRH
                     MSNKGMEHLYNMKCKNVVPLYDLLLEMLDAHRLHAPAARSAAPMEEENRSQLTTASAS
                     SHSLQSFYINSKEEENMQNTL"
     misc_feature    238..645
                     /note="domain B"
     exon            561..751
                     /number=2
     misc_feature    646..894
                     /note="domain C; DNA binding domain"
     exon            752..868
                     /number=3
     exon            869..1204
                     /number=4
     misc_feature    895..1011
                     /note="domain D"
     misc_feature    1012..1764
                     /note="domain E; ligand binding domain"
     exon            1205..1343
                     /number=5
     exon            1344..1477
                     /number=6
     exon            1478..1661
                     /number=7
     exon            1662..>2792
                     /number=8
     misc_feature    1765..1893
                     /note="domain F"
     3'UTR          1897..>2792
     polyA_signal   2737..2743
BASE COUNT      794 a    669 c    615 g    714 t
ORIGIN

  Query Match             11.9%;  Score 149;  DB 52;  Length 2792;
  Best Local Similarity  69.9%;  Pred. No. 5.75e-118;
  Matches  262;  Conservative    0;  Mismatches 113;  Indels    0;  Gaps    0;

Db   1114 ccattcaatgaagcttctatgatgaccctgttgaccaaccttgcagacagagaattggtg 1173
          || ||||  || || || |||||||      ||||||| | || ||||  || |||||
Qy    703 CCCTTCACCGAGGCCTCCATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTA 762

Db   1174 cacatgatcaactgggcaaagagagttccaggatttgtggatttaacactccatgatcag 1233
          ||||||||| |||||| ||||   |||| || ||||||||  | |  ||    || ||
Qy    763 CACATGATCAGCTGGGCCAAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAA 822

Db   1234 gtccatctactggaatgtgcatgattagagtactgatgatcggcttagtgtggcgctcc 1293
          || |  ||  ||||  |     ||| | ||| |   | ||||| ||   |  |||||||||
Qy    823 GTGCGGCTCTTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCA 882

Db   1294 atggaagacccaggaaagctttttgtttgcacctaacctgttactggacaggaatcaaggg 1353
          || || ||||| || |||||   | ||||| ||   | ||   | ||||||||| || | |||
Qy    883 ATTGACCACCCCGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGG 942

Db   1354 aaatgtgtagagggcatggtggaaatctttgacatgctgctggctactgctgctcggttt 1413
          ||||| ||||| || ||  |||||||||||||||||||| ||||| ||| || |  |||||
Qy    943 AAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTT 1002

Db   1414 cgcatgatgaaccttcaaggggaggaatttgtgtgccttaagtccatcatcctgctcaat 1473
          ||    | | || || |||     | |||| | | ||   | ||| |||| ||||||||||||
Qy   1003 CGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAAT 1062

Db   1474 tctggtgtgtacact 1488
          ||  || ||||| ||
Qy   1063 TCCAGTATGTACCCT 1077


RESULT      7
LOCUS       OAESIREC     1966 bp    RNA             MAM       08-OCT-1996
DEFINITION  O.aries mRNA for estrogen receptor.
ACCESSION   Z49257
NID         g1617201
KEYWORDS    estrogen receptor.
SOURCE      sheep.
  ORGANISM  Ovis aries
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Artiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Caprinae; Ovis
REFERENCE   1  (bases 1 to 1966)
  AUTHORS   Madigou,T., Tiffoche,C., Lazennec,G., Pelletier,J. and
            Thieulant,M.L.
  TITLE     The sheep estrogen receptor: cloning and regulation of expression
            in the hypothalamo-pituitary axis
  JOURNAL   Mol. Cell. Endocrinol. 121 (2), 153-163 (1996)
  MEDLINE   97047396
REFERENCE   2  (bases 1 to 1966)
  AUTHORS   Madigou,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-MAY-1995) Thierry Madigou, Interactions Cellulaires
            et Moleculaires, CNRS URA 256, Rennes, 35042, FRANCE
FEATURES             Location/Qualifiers
     source          1..1966
                     /organism="Ovis aries"
                     /strain="Romanov"
                     /clone="lambda P112"
                     /dev_stage="adult"
                     /tissue_type="anterior pituitary gland"
                     /clone_lib="cDNA library"
                     /sex="Female"
     mRNA            1..1966
                     /product="estrogen receptor"
BASE COUNT      439 a    604 c    549 g    374 t
```

```
ORIGIN

  Query Match             11.8%;  Score 148;  DB 47;  Length 1966;
  Best Local Similarity  69.9%;  Pred. No. 6.19e-117;
  Matches  260;  Conservative    0;  Mismatches 112;  Indels    0;  Gaps    0;

Db   1158 cctttcagtgaagcttctatgatgaggcttgctgaccagcttcccgacaggaagctggta 1217
          || ||||  ||||| || ||||||    |  ||||||     | || |||| |||| |||||
Qy    703 CCCTTCACCGAGGCCTCCATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTA 762

Db   1218 cacatgatcaactgggcaaagaggtgccaggatttgtggatcttggccctccagcatcaa 1277
          |||||||||| |||||| |||| | | || || |||||||| ||    |||    ||| |||
Qy    763 CACATGATCAGCTGGGCCAAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAA 822

Db   1278 gtccacctttggaatgtgcctggctagagatcctcatgattggtcttgtctggcgctcc 1337
          || |  || |||||  |     ||| | ||| |   | ||||| || ||   | ||||||||
Qy    823 GTGCGGCTCTTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCA 882

Db   1338 atggagcatccaggaagctcctatttgctcccaaccttctcctggacaggaaccaggga 1397
          || || || || || |||||| | ||||||||    | ||| | |||||||||  |  ||||
Qy    883 ATTGACCACCCCGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGG 942

Db   1398 aaatgtgtggaaggcatggtggagatctttgacatgttgctggctacatcgtctcggttc 1457
          ||||| || ||||| ||   |||| |||||||||||| | ||||| ||   | ||  ||||
Qy    943 AAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTT 1002

Db   1458 cgtatgatgaatctccagggagaagagtttgtgtgcctcaaatccatcatcttgcttaat 1517
          ||    | | || |||||      |||| | | | | ||   ||||  |||| ||| ||| | |||
Qy   1003 CGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAAT 1062

Db   1518 tctggggtgtac 1529
          ||  |  |||||
Qy   1063 TCCAGTATGTAC 1074


RESULT      8
LOCUS       SSOESREC     1662 bp    RNA             VRT       01-AUG-1995
DEFINITION  S.salar mRNA for oestrogen receptor.
ACCESSION   X89959
NID         g929901
KEYWORDS    oestrogen receptor.
SOURCE      Atlantic salmon.
  ORGANISM  Salmo salar
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Actinopterygii; Neopterygii; Teleostei; Euteleostei;
            Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
REFERENCE   1  (bases 1 to 1662)
  AUTHORS   Rogers,S.A., Llewellyn,L., Ramsurn,V.P., Sweeney,G.E. and Wigham,T.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1662)
  AUTHORS   Rogers,S.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JUL-1995) S.A. Rogers, School of Molecular and
            Medical Biosciences, University of Wales, PO Box 911, Cardiff, CF1
            3US, UK
FEATURES             Location/Qualifiers
     source          1..1662
                     /organism="Salmo salar"
                     /dev_stage="adult"
                     /tissue_type="liver"
     CDS             <1..1608
                     /codon_start=1
                     /product="oestrogen receptor"
                     /db_xref="PID:g929902"
                     /translation="SRMLIDPPRIGSMQSLGSSPIGPLVFVSSSPQLSPFLHPPGHHG
                     LPSQSYYLEISSIPLYRSSVVINQLSASEEKLCITSNRQQSYAAAGSGVRVFEMANET
                     RYCAVCSDFASGYHYGFWSCEGCKAFFKRSIQGHNDYMCPATNQCTMDRNRRKSCQAC
                     RLRKCYEVGMVKGGLRKDRGGRVLRKDKRYCGPAGDREKPYGDLEHRTAPPQDGGRNS
                     SSSSLSGGGGWCGPRITMPPPQVLFLLQGAPPALCSRQKVARPYTEVTMMTLLTSMA
                     DKELVHMIAWAKKVPGFQELSLHDQVQLLESSWLEVLMIGLIWRSIHCPGKLIFAQDL
                     ILDRSEGDCVEGMAEIFDMLLATVSRFRMLKLKPEEFVCLKAIILLNSGAFSFCSNSV
                     ESLHNSSAVESMLDNITDALIHHISHSGASVQQQPRRQVQLLLLLSHIRHMSNKGMEH
                     LYSIKCKNKVPLYDLLLEMLDGHRIQSPGKVAQAGEQTEGPSTTTTTSTGSSIGPMRG
                     SQDIHIRSPGSGVLQYGSPSSDQMPIP"
BASE COUNT      383 a    481 c    461 g    336 t      1 others
ORIGIN

  Query Match             11.2%;  Score 140;  DB 52;  Length 1662;
  Best Local Similarity   69.3%;  Pred. No. 1.06e-108;
  Matches  251;  Conservative    0;  Mismatches 111;  Indels    0;  Gaps    0;

Db     784 cccttcacagaggccagcatgatgaccctgctcaccagcatggccgacaaggagctggtg 843
           |||| ||| |||| | ||||||||     || ||||   |||||||||||||| ||||
Qy     703 CCCTTCACTGAGGCCTCCATGATGATGTCCCTGACCAAGTTGGCGGACAAGGAGTTGGTA 762

Db     844 cacatgatcgcttgggctaaggaagtacccaggtttccaggagctgtctctcatgaccag 903
           |||||||||    ||||| |||||   | || || ||    ||||||    ||    |||||
Qy     763 CACATGATCAGCTGGGCCAAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAA 822

Db     904 gtgcagctgctggagagttcctggctaaaatgctaatgatcggactcatatggaggtcc 963
           |||| |||  |||||||| |   ||| |||||||| | ||||| || || || ||| | ||
Qy     823 GTGCGGCTCTTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCA 882

Db     964 atccactgccctgggaaactcatctttgcccaggacctcatactggacaggagtgaaggg 1023
           ||  ||  ||| ||  || |||||||||||| |  || ||  | |||||||||  ||| |||
Qy     883 ATTGACCACCCCGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGG 942

Db    1024 gactgtgtggagggtatggctgagatcttcgacatgctcctggccactgtgtctcacttc 1083
            | || || || || ||     || ||||| |||||||||||||| |||   ||  | ||
Qy     943 AAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTT 1002

Db    1084 cgcatgctcaaactgaagcctgaggatttctctctcctcaaagccatcatcttactcaac 1143
            |     | | |||||  | |    | || | | | || |||| ||||| ||| |||||||
Qy    1003 CGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAAT 1062

Db    1144 tc 1145
           ||
Qy    1063 TC 1064


RESULT      9
LOCUS       SSESTREC      1788 bp    RNA             MAM       15-SEP-1994
DEFINITION  S.scrofa mRNA for estradiol receptor.
ACCESSION   Z37167
NID         g587554
KEYWORDS    estradiol receptor.
SOURCE      pig.
  ORGANISM  Sus scrofa
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Artiodactyla; Suiformes; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 1788)
  AUTHORS   Bokenkamp,D., Jungblut,P.W. and Thole,H.H.
  TITLE     The C-terminal half of the porcine estradiol receptor contains no
            post-translational modification: determination of the primary
            structure
  JOURNAL   Mol. Cell. Endocrinol. 104 (2), 163-172 (1994)
  MEDLINE   95080454
REFERENCE   2  (bases 1 to 1788)
  AUTHORS   Boekenkamp,D.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-SEP-1994) Dirk Boekenkamp, Max-Planck-Institut fuer
            experimentelle Endokrinologie, Feodor-Lynen Strasse 7, Hannover,
            30625, Germany
FEATURES             Location/Qualifiers
     source          1..1788
                     /organism="Sus scrofa"
                     /tissue_type="Uterus"
                     /sex="Female"
     CDS             1..1788
                     /codon_start=1
                     /product="estradiol receptor"
                     /db_xref="PID:g587555"
                     /translation="MTMTLHTKASGMALLHQIQANELEPLNRPQLKIPLERPLGEVYV
                     DSSKPAVYNYPEGAAYDFNAAAAASAPVYGQSGLAYGPGSEAAAFGANGLGGFQPLNS
                     VSPSPLVLLHPPPQLSPFLHPHGQQVPYYLENEPSGYAVREAGPPAFYRPNSDNRRQG
                     GRERLASTSDKGSMAMESAKETRYCAVCNDYASGYHYGVWSCEGCKAFFKRSIQGHND
                     YMCPATNQCTIDKNRRKSCQACRLRKCYEVGMMKGGIRKDRRGGRMLKHKRQRDDGEG
                     RNEAVPPGDMRSANLWPSPLLIKHTKKNSPVLSLTADQMISALLEAEPPIIYSEYDPT
                     RPLSEASMMGLLTNLADRELVHMINWAKRVPGFLDLSLHDQVHLLECAWLEILMIGLV
                     WRSMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEFVCLKS
                     IILLNSGVYTFLSSTLKSLEEKDHIHRVLDKITDTLIHLMAKAGLTLQQQHRRLAQLL
                     LILSHFRHMSNKGMEHLYNMKCKNVVPLYDLLLEMLDAHRLHAPTNLGGPPPEDMSQS
                     QLATSGSTPSHSLQMYYITGEAENFPTTI"
BASE COUNT      410 a    530 c    504 g    344 t
ORIGIN

  Query Match             11.2%;  Score 140;  DB 48;  Length 1788;
  Best Local Similarity   68.8%;  Pred. No. 1.06e-108;
  Matches  255;  Conservative    0;  Mismatches 115;  Indels    0;  Gaps    0;

Db    1006 cccttcagtgaggcttcaatgatgggcttgctgaccaacctggcagacaaggagctggta 1065
           ||| |||  ||||| || ||||||    |  ||||||||  | || |||| |||| ||||
Qy     703 CCCTTCACTGAGGCCTCCATGATGATGTCCCTGACCAAGTTGGCGGACAAGGAGTTGGTA 762

Db    1066 cacatgatcaactgggcaaagagggtgccaggattttggatttaagctcccatgatcaa 1125
           |||||||||| |||||| |||| | | || || ||| ||||  | |||||     || |||
Qy     763 CACATGATCAGCTGGGCCAAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAA 822

Db    1126 gtgcatcttctggaatgtgcctggctagagatcctcatgattggtcttgtctggcactcc 1185
           ||||  ||  ||||  |     ||| | ||| |   | ||||| || ||   | ||||||||
Qy     823 GTGCGGCTCTTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCA 882

Db    1186 atggagcacccagggaagctcctgtttgctcctaacttgctcctggacaggaatcaaggc 1245
           || || ||||| || |||||| | ||||||||    |  |   ||||||||     | |
Qy     883 ATTGACCACCCCGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGG 942

Db    1246 aagtgtgtcgagggaatggtggagatctttgacatgttgctggctacatcatctcgcttc 1305
           || || || || ||||| |||| |||||||||||| | ||||| ||  | ||  | ||
Qy     943 AAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTT 1002

Db    1306 cgtatgatgaatctccagggagaggagtttgtgtgcctcaaatccatcattttacttaat 1365
           ||   | | || |||||     | || | | | ||  ||||  |||| ||  |||| |||
Qy    1003 CGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAAT 1062

Db    1366 tctggagtgtac 1377
           ||  |  |||||
Qy    1063 TCCAGTATGTAC 1074


RESULT     10
LOCUS       RNESTROR      1880 bp    RNA             ROD       29-OCT-1996
DEFINITION  Rat mRNA for estrogen receptor.
ACCESSION   X61098
NID         g56120
KEYWORDS    estrogen receptor; steroid hormone receptor.
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 1880)
  AUTHORS   Maggi,A.M.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-JUN-1991) A.M.A. Maggi, Milano Mol Pharm Lab, Inst
            Pharmacological Sciences, Piazza Durante 11, 20131 Milano, ITALY
REFERENCE   2  (bases 1 to 1880)
  AUTHORS   Spreafico,E., Bettini,E., Pollio,G. and Maggi,A.
  TITLE     Nucleotide sequence of estrogen receptor cDNA from Sprague-Dawley
            rat
  JOURNAL   Eur. J. Pharmacol. 227 (3), 353-356 (1992)
  MEDLINE   93114361
FEATURES             Location/Qualifiers
     source          1..1880
                     /organism="Rattus norvegicus"
                     /strain="Sprague-Dawley"
                     /tissue_type="uterus"
     mRNA            1..1880
                     /evidence=experimental
                     /note="estrogen receptor"
     CDS             74..1876
                     /codon_start=1
                     /product="estrogen receptor"
                     /db_xref="PID:g56121"
                     /db_xref="SWISS-PROT:P06211"
                     /translation="MIMILHIKASGMALLHQIQGNELEPLNRPQLKMPMERALGEVYV
                     DNSKPAVFNYPEGAAYEFNAAAAAAAAGASAPVYGQSSITYGPGSEAAAFGANSLGAF
                     PQLNSVSPSPLMLLHPPPHVSPFLHPHGHQVPYYLENEPSAYAVRDTGPPAFYRSNSD
                     NRRQNGRERLSSSSEKGNMIMESAKETRYCAVCNDYASGYHYGVWSCEGCKAFFKRSI
                     QGHNDYMCPAINQCTIDKNRRKSCQACRLRKCYEVGMMKGGIRKDRRGGRMLKHKRQR
                     DDLEGRNEMGTSGDMRAANLWPSPLVIKHTKKNSPALSLTADQMVSALLDAEPPLIYS
                     EYDPSRPFSEASMMGLLTNLADRELVHMINWAKRVPGFGDLNLHDQVHLLECAWLEIL
                     MIGLVWRSMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEF
                     VCLKSIILLNSGVYTFLSSTLKSLEEKDHIHRVLDKITDTLIHLMAKAGLTLQQQHRR
                     LAQLLLILSHIRHMSNKGMEHLYNMKCKNVVPLYDLLLEMLDAHRLHAPASRMGVPPE
                     EPSQSQLTTTSSTSAHSLQTYYIPPEAEGFPNTI"
BASE COUNT      438 a    548 c    505 g    389 t
ORIGIN

  Query Match             11.2%;  Score 140;  DB 91;  Length 1880;
  Best Local Similarity   82.4%;  Pred. No. 1.06e-108;
  Matches  178;  Conservative    0;  Mismatches   38;  Indels    0;  Gaps    0;

Db     628 ggagactcgctactgtgctgtgtgtaacgactatgcctctggctaccattatggggtctg 687
           |||  ||| || ||| ||||| || | ||| || || || || || ||||| |||||
Qy     273 GGATGCTCACTTCTGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTG 332

Db     688 gtcctgtgaaggctgcaaggctttctttaagagaagcattcaaggacataatgactacat 747
           ||| |||||||| || ||||| || ||||| ||||||||||||||||||||||| || ||
Qy     333 GTCGTGTGAAGGATGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATAT 392

Db     748 atgtccagctacaaaccaatgcaccatcgataagaaccggaggaagagttgccaggcctg 807
            |||||||||||||| || || || |||||||| |||||| | ||||| |||||||||||
Qy     393 TTGTCCAGCTACAAATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTG 452

Db     808 ccggctgcgcaagtgttacgaagtgggcatgatgaa 843
           ||| || || ||||||||||||||||| ||| ||||
Qy     453 CCGACTTCGGAAGTGTTACGAAGTGGGAATGGTGAA 488


RESULT    11
LOCUS       RNERR        2090 bp    RNA             ROD       12-SEP-1993
DEFINITION  Rat mRNA for estrogen receptor
ACCESSION   Y00102
NID         g56110
KEYWORDS    estrogen receptor.
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 2090)
  AUTHORS   Muramatsu,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-MAR-1987) Masami Muramatsu,Dept of Biochemistry,
            Faculty of Medicine The University of Tokyo, Hongo 7-3-1,
            Bunkyo-Ku, Tokyo 113 Japan
REFERENCE   2  (bases 1 to 2090)
  AUTHORS   Koike,S., Sakai,M. and Muramatsu,M.
  TITLE     Molecular cloning and characterization of rat estrogen receptor
            cDNA
  JOURNAL   Nucleic Acids Res. 15 (6), 2499-2513 (1987)
  MEDLINE   87174780
FEATURES             Location/Qualifiers
     source          1..2090
                     /organism="Rattus norvegicus"
                     /strain="Wistar"
                     /tissue_type="utems"
                     /clone_lib="rat uterus cDNA 1 "
     CDS             210..2012
                     /note="estrogen receptor protein (AA 1-600)"
                     /codon_start=1
                     /db_xref="PID:g56111"
```

```
                     /db_xref="SWISS-PROT:P06211"
                     /translation="MIMILHIKASGMALLHQIQGNELEPLNRPQLKMPMERALGEVYV
                     DNSKPAVFNYPEGAAYEFNAAAAAAAAGASAPVYGQSSITYGPGSEAAAFGANSLGAF
                     PQLNSVSPSPLMLLHPPPHVSPFLHPHGHQVPYYLENEPSAYAVRDTGPPAFYRSNSD
                     NRRQNGRERLSSSSEKGNMIMESAKETRYCAVCNDYASGYHYGVWSCEGCKAFFKRSI
                     QGHNDYMCPAINQCTIDKNRRKSCQACRLRKCYEVGMMKGGIRKDRRGGRMLKHKRQR
                     DDLEGRNEMGTSGDMRAANLWPSPLVIKHTKKNSPALSLTADQMVSALLDAEPPLIYS
                     EYDPSRPFSEASMMGLLTNLADRELVHMINWAKRVPGFGDLNLHDQVHLLECAWLEIL
                     MIGLVWRSMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEF
                     VCLKSIILLNSGVYTFLSSTLKSLEEKDHIHRVLDKINDTLIHLMAKAGLTLQQQHRR
                     LAQLLLILSHIRHMSNKGMEHLYNMKCKNVVPLYDLLLEMLDAHRLHAPASRMGVPPE
                     EPSQSQLTTTSSTSAHSLQTYYIPPEAEGFPNTI"
BASE COUNT      488 a    619 c    537 g    446 t
ORIGIN

  Query Match             11.2%;  Score 140;  DB 91;  Length 2090;
  Best Local Similarity   82.4%;  Pred. No. 1.06e-108;
  Matches  178;  Conservative    0;  Mismatches   38;  Indels    0;  Gaps    0;

Db     764 ggagactcgctactgtgctgtgtgtaacgactatgcctctggctaccattatggggtctg 823
           |||  ||| || ||| ||||| || | ||| || || || || || ||||| |||||
Qy     273 GGATGCTCACTTCTGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTG 332

Db     824 gtcctgtgaaggctgcaaggctttctttaagagaagcattcaaggacataatgactacat 883
           ||| |||||||| || ||||| || ||||| ||||||||||||||||||||||| || ||
Qy     333 GTCGTGTGAAGGATGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATAT 392

Db     884 atgtccagctacaaaccaatgcaccatcgataagaaccggaggaagagttgccaggcctg 943
            |||||||||||||| || || || |||||||| |||||| | ||||| |||||||||||
Qy     393 TTGTCCAGCTACAAATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTG 452

Db     944 ccggctgcgcaagtgttacgaagtgggcatgatgaa 979
           ||| || || ||||||||||||||||| ||| ||||
Qy     453 CCGACTTCGGAAGTGTTACGAAGTGGGAATGGTGAA 488


RESULT    12
LOCUS       S52128       2090 bp    mRNA            ROD       12-MAR-1993
DEFINITION  estrogen receptor [rats, brain, mRNA, 2090 nt]
ACCESSION   S52128
NID         g262708
KEYWORDS    .
SOURCE      Rattus sp. brain.
  ORGANISM  Rattus sp.
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 2090)
  AUTHORS   Bettini,E., Pollio,G., Santagati,S. and Maggi,A.
  TITLE     Estrogen receptor in rat brain: presence in the hippocampal
            formation
  JOURNAL   Neuroendocrinology 56 (4), 502-508 (1992)
  MEDLINE   93116840
  REMARK    GenBank staff at the National Library of Medicine created this
            entry [NCBI gibbsq 121564] from the original journal article.
            This sequence comes from Fig. 1.
FEATURES             Location/Qualifiers
     source          1..2090
                     /organism="Rattus sp."
BASE COUNT      488 a    619 c    537 g    446 t
ORIGIN

  Query Match             11.2%;  Score 140;  DB 92;  Length 2090;
  Best Local Similarity   82.4%;  Pred. No. 1.06e-108;
  Matches  178;  Conservative    0;  Mismatches   38;  Indels    0;  Gaps    0;

Db     764 ggagactcgctactgtgctgtgtgtaacgactatgcctctggctaccattatggggtctg 823
           |||  ||| || ||| ||||| || | ||| || || || || || ||||| |||||
Qy     273 GGATGCTCACTTCTGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTG 332

Db     824 gtcctgtgaaggctgcaaggctttctttaagagaagcattcaaggacataatgactacat 883
           ||| |||||||| || ||||| || ||||| ||||||||||||||||||||||| || ||
Qy     333 GTCGTGTGAAGGATGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATAT 392
```

```
Db    884 gtgtccagccaccaaccagtgcaccattgataaaaacaggaggaagagctgccaggcctg 943
           |||||||| || || ||||| || || ||||||||| || | |||||||||||||||||
Qy    393 TTGTCCAGCTACAAATCAGTGTACAATCGATAAAAACTGGCGCAAGAGCTGCCAGGCCTG 452

Db    944 ccggctccgcaaatgctacgaagtgggaatgatgaa 979
          ||| || || || || ||||||||||||||| ||||
Qy    453 CCGACTTCGGAAGTGTTACGAAGTGGGAATGGTGAA 488


RESULT   13
LOCUS       HSU47678     2106 bp    mRNA            PRI       17-MAY-1996
DEFINITION  Human 80 kDa estrogen receptor mRNA, containing exon 6 and 7
            duplication, complete cds.
ACCESSION   U47678
NID         g1197854
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2106)
  AUTHORS   Pink,J.J., Wu,S.Q., Wolf,D.M., Bilimoria,M.M. and Jordan,V.C.
  TITLE     A novel 80 kDa human estrogen receptor containing a duplication of
            exons 6 and 7
  JOURNAL   Nucleic Acids Res. 24 (5), 962-969 (1996)
  MEDLINE   96174665
REFERENCE   2  (bases 1 to 2106)
  AUTHORS   Jordan,C.V.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-JAN-1996) Craig V. Jordan, Robert H. Lurie Cancer
            Center, Northwestern University Medical Center, 303 E. Chicago
            Ave., Olson 8258, Chicago, IL 60611, USA
FEATURES             Location/Qualifiers
     source          1..2106
                     /organism="Homo sapiens"
                     /chromosome="6"
                     /map="6q25"
                     /cell_line="MCF-7"
                     /tissue_type="breast"
     exon            1..450
                     /number=1
     CDS             1..2106
                     /note="80 kDa estrogen receptor with an exon 6 and 7
                     repeat; It is able to bind estrogen"
                     /codon_start=1
                     /product="80 kDa estrogen receptor"
                     /db_xref="PID:g1197855"
                     /translation="MTMTLHTKASGMALLHQIQGNELEPLNRPQLKIPLERPLGEVYL
                     DSSKPAVYNYPEGAAYEFNAAAAANAQVYGQTGLPYGPGSEAAAFGSNGLGGFPPLNS
                     VSPSPLMLLHPPPQLSPFLQPHGQQVPYYLENEPSGYTVREAGPPAFYRPNSDNRRQG
                     GRERLASTNDKGSMAMESAKETRYCAVCNDYASGYHYGVWSCEGCKAFFKRSIQGHND
                     YMCPATNQCTIDKNRRKSCQACRLRKCYEVGMMKGGIRKDRRGGRMLKHKRQRDDGEG
                     RGEVGSAGDMRAANLWPSPLMIKRSKKNSLALSLTADQMVSALLDAEPPILYSEYDPT
                     RPFSEASMMGLLTNLADRELVHMINWAKRVPGFVDLTLHDQVHLLECAWLEILMIGLV
                     WRSMEHPVKLLFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEFVCLKS
                     IILLNSGVYTFLSSTLKSLEEKDHIHRVLDKITDTLIHLMAKAGLTLQQQHQRLAQLL
                     LILSHIRHMRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEFVCLKSIILLNSGVYT
                     FLSSTLKSLEEKDHIHRVLDKITDTLIHLMAKAGLTLQQQHQRLAQLLLILSHIRHMS
                     NKGMEHLYSMKCKNVVPLYDLLLEMLDAHRLHAPTSRGGASVEETDQSHLATAGSTSS
                     HSLQKYYITGEAEGFPATV"
     exon            451..642
                     /number=2
     exon            643..759
                     /number=3
     exon            760..1095
                     /number=4
     exon            1096..1233
                     /number=5
     exon            1234..1368
                     /number=6
     exon            1369..1551
                     /number=7
     exon            1551..1685
                     /note="exon 6'"
     exon            1686..1868
                     /note="exon 7'"
     exon            1869..2106
                     /number=8
BASE COUNT      496 a    602 c    587 g    421 t
ORIGIN

  Query Match              10.9%;  Score 136;  DB 76;  Length 2106;
  Best Local Similarity  63.8%;  Pred. No. 1.35e-104;
  Matches  350;  Conservative    0;  Mismatches  196;  Indels    3;  Gaps    2;

Db   1006 cccttcagtgaagcttcgatgatgggcttactgaccaacctggcagacagggagctggtt 1065
          |||||||  || || || ||||||   |  ||||||||  |||| |||| |||| |||| 
Qy    703 CCCTTCACCGAGGCCTCCATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTA 762

Db   1066 cacatgatcaactgggcgaagagggtgccaggctttgtggatttgaccctccatgatcag 1125
          |||||||||| |||||| |||| | | || || ||||||||  | | ||| |  || || 
Qy    763 CACATGATCAGCTGGGCCAAGAAGATCCCCGGATTTGTGGAGCTCAGCCTGCTGGACCAA 822

Db   1126 gtccaccttctagaatgtgcctggctagagatcctgatgattggtctcgtctggcgctcc 1185
          || | |||  | ||  |    ||| | ||| |  | ||||| || ||  | |||||||| 
Qy    823 GTGCGCCTCTTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCA 882

Db   1186 atggagcacccagtgaagctactgtttgctcctaacttgctcttggacaggaaccaggga 1245
          || || ||||| |  |||||  | ||||||||  |  |  |  |||||||| | ||||| 
Qy    883 ATTGACCACCCGGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATCAGGGG 942

Db   1246 aaatgtgtagagggcatggtggagatcttcgacatgctgctggctacatcatctcggttc 1305
          ||||| ||||| || ||  |||| ||||| |||||||| ||||| || |  ||  |||| 
Qy    943 AAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTT 1002

Db   1306 cgcatgatgaatctgcagggagaggagtttgtgtgcctcaaatctattattttgcttaat 1365
          ||   | | || || ||     | || |   | ||  ||||  |  | ||  |||| |||
Qy   1003 CGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAAT 1062

Db   1366 tctggagtgtacacatttctgtccagcaccctgaagtctctggaagagaaggaccatatc 1425
          ||  |  ||||| |  |       || |||| | |  ||            ||   |  |
Qy   1063 TCCAGTATGTACCCTCTGGTCACAGCGACCCAGGATGCTGAC-AGCAGCCGGAAGCTGGC 1121

Db   1426 caccgagtcctggacaagatcacagacactttgatccacctgatggccaaggcaggcctg 1485
             |  |  |||  |      || ||   |||| |      ||| ||||||   ||| | 
Qy   1122 TCACTTG--CTGAACGCCGTGACCGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATC 1179

Db   1486 accctgcagcagcagcaccagcggctggcccagctcctcctcatcctctcccacatcagg 1545
           ||    |||||||   |   ||||||||  |||||||  |  |||||||||||||||||
Qy   1180 TCCTCCTAGCAGCAATCCATCCGGCTGGCTAAGCTCCTGATGCTCCTCTCCCACATCAGG 1239

Db   1546 cacatgagg 1554
          ||   ||||
Qy   1240 CATGCGAGG 1248


RESULT   14
LOCUS       SMOERA       1278 bp    mRNA            VRT       15-JUN-1990
DEFINITION  Rainbow trout estrogen receptor mRNA, 3' end.
ACCESSION   M31559
NID         g213783
KEYWORDS    estrogen receptor.
SOURCE      Rainbow trout (female) liver, cDNA to mRNA.
  ORGANISM  Oncorhynchus mykiss
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Actinopterygii; Neopterygii; Teleostei; Euteleostei;
            Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
REFERENCE   1  (bases 1 to 1278)
  AUTHORS   Pakdel,F., Le Guellec,C., Vaillant,C., Le Roux,M.G. and
            Valotaire,Y.
  TITLE     Identification and estrogen induction of two estrogen receptors
            (ER) messenger ribonucleic acids in the rainbow trout liver:
            Sequence homology with other ERs
  JOURNAL   Mol. Endocrinol. 3, 44-51 (1989)
  MEDLINE   89127284
```

```
FEATURES             Location/Qualifiers
     source          1..1278
                     /organism="Oncorhynchus mykiss"
     CDS             <1..1278
                     /note="estrogen receptor"
                     /codon_start=1
                     /db_xref="PID g213784"
                     /translation="CSDFASGYHYGVWSCEGCKAFFKRSIQGHNDYMCPATNQCTMDR
                     NRRKSCQACRLRKCYEVGMVKGGLRKDRGGRVLRKDKRYCGPAGDREKPTVTWSTGQR
                     PQDGGRNSSSSLNGGGGWRGPRITMPPEQVLFLLQGQTPALCSRQKVARPYTEVTMMT
                     LLTSMADKELVHMIAWAKKVPGFQELSLHDQVQLLESSWLEVLMIGLIWRSIHCPGKL
                     IFAQDLILDRSEGDCVEGMAEIFDMLLATVSRFGMLKLKPEEFVCLKAIILLNPGAFS
                     FCSNSVESLHNSSAVESMLDNITDALIHHISHSGASVQQQPRRQAQLLLLLSHIRHMS
                     NKGMEHLYSIKCKNKVPLYDLLLEMLDGHRLQAPGKVAQAGEQTEGPSTTTTTSTGSS
                     IGPMRGSQDTHIRSPGSGVLQYGSPSSDQMPIP"
BASE COUNT      294 a    369 c    366 g    249 t
ORIGIN

  Query Match             10.8%;  Score 135;  DB 52;  Length 1278;
  Best Local Similarity   68.8%;  Pred. No. 1.43e-103;
  Matches  247;  Conservative    0;  Mismatches 112;  Indels    0;  Gaps    0;

Db     454 ccctacacagaggtcaccatgatgaccctgctcaccagcatggctgacaaggagctggtg 513
           |||| ||| |||| | |||||||||      || ||||    |||| ||||||||| ||||
Qy     703 CCCTTCACCGAGGCCTCCATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTA 762

Db     514 cacatgatcgcttgggctaagaaagtaccaggtttccaggagctgtctctccatgaccag 573
           |||||||||    ||||| |||||  | || || ||   ||||||     ||    |||||
Qy     763 CACATGATCAGCTGGGCCAAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAA 822

Db     574 gtgcagctgctggagagttcctggctggaggtgctgatgatcggactcatatggcggtcc 633
           |||| |||  ||||||| |   ||| |||||||| | ||||| || || || ||||| ||
Qy     823 GTGCGGCTCTTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCA 882

Db     634 atccactgccctgggaaactcatcttcgcccaggacctcatactggacaggagtgaaggg 693
           ||   ||   ||| || || |||||||| || |  || ||   | |||||||||   ||| |||
Qy     883 ATTGACCACCCCGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGG 942

Db     694 gactgtgtggagggtatggctgagatcttcgacatgctcctggccactgtgtctcgcttc 753
            | || || || || ||     || ||||| ||||||||||||| |||    ||   | ||
Qy     943 AAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTT 1002

Db     754 ggcatgcttaaactgaagcctgaggagtttgtgtgcctcaaagccatcatcttgctcaa 812
            |    | | ||||| | |     | || | | | ||   |||| ||||| ||| |||||||
Qy    1003 CGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAA 1061


RESULT    15
LOCUS        I08538       2092 bp                     PAT       14-NOV-1994
DEFINITION  Sequence 1 from Patent WO 8705049.
ACCESSION   I08538
NID         g588753
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2092)
  AUTHORS   Shine,J.
  TITLE     EUCARYOTIC EXPRESSION OF STEROID RECEPTOR PROTEINS
  JOURNAL   Patent: WO 8705049-A 1 27-AUG-1987;

FEATURES             Location/Qualifiers
     source          1..2092
                     /organism="unknown"
BASE COUNT      473 a    605 c    593 g    421 t
ORIGIN

  Query Match             10.8%;  Score 135;  DB 55;  Length 2092;
  Best Local Similarity   77.8%;  Pred. No. 1.43e-103;
  Matches  189;  Conservative    0;  Mismatches  54;  Indels    0;  Gaps    0;

Db     832 ggagactcgctactgtgcagtgtgcaatgactatgcttcaggctaccattatggagtctg 891
           |||  ||| || ||| || || ||||  || || || || || || || |||||||||||
Qy     273 GGATGCTGACTTCTGCCTGTCTGCAGCGATTACGCATGGGATATCACTATGGAGTCTG 332

Db     892 gtcctgtgagggctgcaaggccttcttcaagagaagtattcaaggacataacgactatat 951
           |||  ||||| || || |||||||| || || ||||| |||||||||||||| || |||||
Qy     333 GTCGTGTGAAGGATGTAAGGCCTTTTTAAAAGAAGCATTCAAGGACATAATGATTATAT 392

Db     952 gtgtccagccaccaaccagtgcaccattgataaaaacaggaggaagagctgccaggcctg 1011
            ||||||| || || ||||| || || ||||||||| || | ||||||||||||||||||
Qy     393 TTGTCCAGCTACAAATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTG 452

Db    1012 cggctccgcaaatgctacgaagtgggaatgatgaaaggtgggatacgaaaagaccgaag 1071
           ||| || || || || ||||||||||||||| ||||  ||||      || | |||  || |
Qy     453 CCGACTTCGGAAGTGTTACGAAGTGGGAATGGTGAAGTGTGGCTCCCGGAGAGAGAGATG 512

Db    1072 agg 1074
            ||
Qy     513 TGG 515

Search completed: Thu Dec 18 18:53:53 1997
Job time : 1074 secs.
```

```
***************************************************************************
   _    _   ___     ___    ___    ___    _   _
  | \  / | |   \   /       |   \  /      | | | |
  |  \/  | |___/   \___    |___/  |      | |_| |
  |      | |           \   |  \   |      |  _  |
  |      | |        ___/   |   \  \___   | | | |  (TM)
***************************************************************************

          Release 2.1D John F. Collins, Biocomputing Research Unit.
          Copyright (c) 1993, 1994, 1995  University of Edinburgh, U.K.
                  Distribution rights by IntelliGenetics, Inc.

MPsrch_nn   n.a. - n.a. database search, using Smith-Waterman algorithm

Run on:          Thu Dec 18 20:20:19 1997   MasPar time 1584.91 Seconds
                                            1315.235 Million cell updates/sec
Tabular output not generated

Title:           >US-08-826-361A-24
Description:     (1-1898) from US08826361A.seq
Perfect Score:   1898
N.A. Sequence.      1 CACGAATCTTTGAGAACATT           TGAATCGAGGGTTTCCACAC 1898
     Comp:            GTGCTTAGAAACTCTTGTAA           ACTTAGCTCCCAAAGGTGTG

Scoring table:   TABLE default
                 Gap 6

Nmatch   STD :   Dbase 0;  Query 0

Searched:        362067 seqs, 549128275 bases x 2

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database:        embl-new3
                   1:BCT 2:FUN 3:GEN1 4:GEN2 5:HTG1 6:HTG2 7:HUM 8:INV
                   9:ORG 10:MAM 11:VRT 12:PLN 13:PRO 14:ROD 15:SYN 16:UNC
                   17:VIR
Database:        genbank99
                   18:BCT1 19:BCT2 20:BCT3 21:BCT4 22:BCT5 23:BCT6 24:BCT7
                   25:BCT8 26:BCT9 27:BCT10 28:BCT11 29:GEN1 30:GEN2
                   31:GEN3 32:HTG1 33:HTG2 34:HTG3 35:INV1 36:INV2 37:INV3
                   38:INV4 39:INV5 40:INV6 41:INV7 42:INV8 43:INV9 44:INV10
                   45:INV11 46:MAM1 47:MAM2 48:MAM3 49:VRT1 50:VRT2 51:VRT3
                   52:VRT4 53:PAT1 54:PAT2 55:PAT3 56:PAT4 57:PAT5 58:PHG
                   59:PLN1 60:PLN2 61:PLN3 62:PLN4 63:PLN5 64:PLN6 65:PLN7
                   66:PLN8 67:PLN9 68:PLN10 69:PLN11 70:PRI1 71:PRI2
                   72:PRI3 73:PRI4 74:PRI5 75:PRI6 76:PRI7 77:PRI8 78:PRI9
                   79:PRI10 80:PRI11 81:PRI12 82:PRI13 83:PRI14 84:PRI15
                   85:ROD1 86:ROD2 87:ROD3 88:ROD4 89:ROD5 90:ROD6 91:ROD7
                   92:ROD8 93:STS 94:SYN 95:UNA 96:VRL1 97:VRL2 98:VRL3
                   99:VRL4 100:VRL5 101:VRL6 102:VRL7 103:VRL8 104:VRL9
                   105:VRL10
Database:        genbank-new3
                   106:BCT 107:GEN1 108:GEN2 109:HTG1 110:HTG2 111:INV
                   112:MAM 113:VRT 114:PHG 115:PLN 116:PRI1 117:PRI2
                   118:ROD 119:SYN 120:UNA 121:VRL
Database:        u-embl50_99
                   122:part1

Statistics:      Mean 11.729;  Variance 4.456;  scale 2.632

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                  SUMMARIES
   *
```

```
Result          Query
  No.   Score   Match Length DB  ID           Description              Pred. No.
--------------------------------------------------------------------------------
    1    1558    82.1   1560 76  HSRNAERB     H.sapiens mRNA for es    0.00e+00
    2    1138    60.0   2555 92  RNU57439     Rattus norvegicus est    0.00e+00
    3     277    14.6    332 70  CJERB        C.jacchus mRNA for es    3.41e-258
    4     217    11.4   1728 51  OKDMER       Killifish mRNA for es    8.99e-193
    5     203    10.7   1798 49  SSESTREC     S.scrofa mRNA for est    1.24e-177
    6     201    10.6   1966 47  OAESTREC     O.aries mRNA for estr    1.78e-175
    7     201    10.6   2095 50  GGERR        Chicken mRNA for estr    1.78e-175
    8     197    10.4   2092 55  I08539       Sequence 1 from Paten    3.65e-171
    9     197    10.4   2092 80  HUMERMCF     Human estrogen recept    3.65e-171
   10     197    10.4   4963 56  I15368       Sequence 16 from pate    3.65e-171
   11     197    10.4   6450 74  HSERR        Human mRNA for oestro    3.65e-171
   12     189    10.0   2792 52  ZFIER        Poephila guttata estr    1.48e-162
   13     184     9.7   1880 91  RNESTROR     Rat mRNA for estrogen    3.46e-157
   14     184     9.7   2090 91  RNERR        Rat mRNA for estrogen    3.46e-157
   15     184     9.7   2090 92  S52129       estrogen receptor [ra    3.46e-157
   16     178     9.4   2078 88  MUSESTRU     Mouse estrogen recept    9.29e-151
   17     175     9.2    586 73  HS275126     H.sapiens mRNA for es    1.50e-147
   18     172     9.1   1662 52  SSOESREC     S.salar mRNA for oest    2.41e-144
   19     170     9.0   1278 52  SMOERA       Rainbow trout estroge    7.23e-142
   20     141     7.4    923 51  S79923       estrogen receptor=rep    1.76e-111
   21     140     7.4   2175 78  HSU47678     Human 80 KDa estrogen    1.99e-110
   22     136     7.2   2449 52  XELESTRREC   Xenopus laevis estrog    3.12e-106
   23     136     7.2   2479 52  XELESTRO     Xenopus laevis estrog    3.12e-106
   24     136     7.2   2573 52  XELER3       Xenopus laevis estrog    3.12e-106
   25     135     7.1    477 46  BTU64962     Bos taurus estrogen r    3.48e-105
   26     133     7.0   2619 90  HUMERVBCA    Human estrogen recept    4.31e-103
   27     109     5.7    225 47  OAU30299     Ovis aries estrogen r    3.13e-78
   28      78     4.1    323 74  HSEST        H. sapiens mRNA for c    4.07e-47
   29      78     4.1    323 117  HSEST        H. sapiens mRNA for c   4.07e-47
   30      77     4.1    400 46  BTEREC       B.taurus mRNA for est    3.85e-46
   31      74     3.9    365 51  OAESREXN4    O.aureus estrogen rec    3.18e-43
c  32      71     3.7   3460 57  I25343       Sequence 1 from paten    2.50e-40
   33      68     3.6    363 84  S71040       estradiol receptor [M    1.87e-37
   34      68     3.6   2153 55  I08706       Sequence 13 from Pate    1.87e-37
   35      68     3.6   2153 76  HSSTHOR2     Human mRNA for steroi    1.87e-37
   36      66     3.5   1344 92  S82459       mERR-2=estrogen respo    1.49e-35
   37      66     3.5   1344 122  MMERR2GEN    M.musculus mRNA for e   1.49e-35
   38      64     3.4   2265 85  MMGRLA       Mouse mRNA for glucoc    1.15e-33
   39      62     3.3   2376 10  ANU87952     Aotus nancymaae gluco    8.69e-32
   40      62     3.3   2376 10  SOU87953     Saguinus oedipus gluc    8.69e-32
   41      62     3.3   2376 116  SOU87953     Saguinus oedipus gluc   8.69e-32
   42      62     3.3   2376 116  ANU87952     Aotus nancymaae gluco   8.69e-32
   43      62     3.3   2575 85  MMGCRR       Mouse mRNA for glucoc    8.69e-32
   44      61     3.2   2391 49  RABPGR       Rabbit progesterone r    1.06e-31
   45      61     3.2   6244 56  I15367       Sequence 15 from pate    1.06e-31


                                  ALIGNMENTS

RESULT      1
LOCUS       HSRNAERB     1560 bp    RNA              PRI       05-SEP-1996
DEFINITION  H.sapiens mRNA for estrogen receptor.
ACCESSION   X99101
NID         g1518262
KEYWORDS    estrogen receptor.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1560)
  AUTHORS   Mosselman,S., Polman,J. and Dijkema,R.
  TITLE     ER beta: identification and characterization of a novel human
            estrogen receptor
  JOURNAL   FEBS Lett. 392 (1), 49-53 (1996)
  MEDLINE   96354875
REFERENCE   2  (bases 1 to 1560)
  AUTHORS   Mosselman,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-JUL-1996) S. Mosselman, N.V.Organon, Biotechnology
            and Biochemistry, PO box 20 Oss, Molenstraat 110 Oss, 5340 BH,
            NETHERLANDS
FEATURES             Location/Qualifiers
     source          1..1560
                     /organism="Homo sapiens"
                     /tissue_type="testis"
                     /dev_stage="adult"
                     /clone_lib="Clontech; HL1010B"
                     /clone_lib="Clontech; 7414-1"
     CDS             19..1452
                     /codon_start=1
                     /product="estrogen receptor beta"
                     /db_xref="PID:e255390"
                     /db_xref="PID:g1718263"
                     /translation="MNYSIPSNVTNLEGGPGRQTISPNVLWPTPGHLSPLVVHRQLSH
                     LYAEPQKSPWCEARSLEHTLPVNRETLKRKVSGNRCASPVTGPGSKRDAHFCAVCSDY
                     ASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRKCYEVG
                     MVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQLVLT
                     LIFAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFDQVR
                     LLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSRF
                     RELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKSG
                     ISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNAHVLRGC
                     KSSITGSECSPAEDSKSKEGSQNPQSQ"
BASE COUNT      367 a    408 c    445 g    340 t
ORIGIN

  Query Match            82.1%;  Score 1558;  DB 76;  Length 1560;
  Best Local Similarity  100.0%;  Pred. No. 0.00e+00;
  Matches 1558;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Db          3 ctatagccctgctgtgatgaattacagcattcccagcaatgtcactaacttggaaggtgg 62
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        220 CTATAGCCCTGCTGTGATGAATTACAGCATTCCCAGCAATGTCACTAACTTGGAAGGTGG 279

Db         63 gcctggtcggcagaccacaagcccaaatgtgttgtggccaacacctgggcacctttctcc 122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        280 GCCTGGTCGGCAGACCACAAGCCCAAATGTGTTGTGGCCAACACCTGGGCACCTTTCTCC 339

Db        123 tttagtggtccatcgccagttatcacatctgtatgcggaacctcaaaagagtccctggtg 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        340 TTTAGTGGTCCATCGCCAGTTATCACATCTGTATGCGGAACCTCAAAAGAGTCCCTGGTG 399

Db        183 tgaagcaagatcgctagaacacaccttacctgtaaacagagagacactgaaaaggaaggt 242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        400 TGAAGCAAGATCGCTAGAACACACCTTACCTGTAAACAGAGAGACACTGAAAAGGAAGGT 459

Db        243 tagtgggaaccgttgcgccagccctgttactggtccaggttcaaagagggatgctcactt 302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        460 TAGTGGGAACCGTTGCGCCAGCCCTGTTACTGGTCCAGGTTCAAAGAGGGATGCTCACTT 519

Db        303 ctgcgctgtctgcagcgattacgcatcgggatatcactatggagtctggtcgtgtgaagg 362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        520 CTGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGG 579

Db        363 atgtaaggccttttttaaaagaagcattcaaggacataatgattatatttgtccagctac 422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        580 ATGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTAC 639

Db        423 aaatcagtgtacaatcgataaaaaccggcgcaagagctgccaggcctgccgacttcggaa 482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        640 AAATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAA 699

Db        483 gtgttacgaagtgggaatggtgaagtgtggctcccggagagagagatgtgggtaccgcct 542
              ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Qy        700 GTGTTACGAAGTGGGAATGGTGAAGTGTGGCTCCCGGAGAGAGAGATGTGGGTACCGCCT 759

Db        543 tgtgcggagacagagaagtgccgacgagcagctgcactgtgccggcaaggccaagagaag 602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        760 TGTGCGGAGACAGAGAAGTGCCGACGAGCAGCTGCACTGTGCCGGCAAGGCCAAGAGAAG 819

Db        603 tggcggccacgcgccccgagtgcgggagctgctgctggacgccctgagccccgagcagct 662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        820 TGGCGGCCACGCGCCCCGAGTGCGGGAGCTGCTGCTGGACGCCCTGAGCCCCGAGCAGCT 879

Db        663 agtgctcaccctcctggaggctgagccgccccatgtgctgatcagccgccccagtgcgcc 722
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        880 AGTGCTCACCCTCCTGGAGGCTGAGCCGCCCCATGTGCTGATCAGCCGCCCCAGTGCGCC 939

Db        723 cttcaccgaggcctccatgatgatgtccctgaccaagttggccgacaaggagttggtaca 782
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        940 CTTCACCGAGGCCTCCATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTACA 999

Db        783 catgatcagctgggccaagaagattcccggctttgtggagctcagcctgttcgaccaagt 842
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1000 CATGATCAGCTGGGCCAAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAAGT 1059

Db        843 gcggctcttggagagctgttggatggaggtgttaatgatggggctgatgtggcgctcaat 902
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1060 GCGGCTCTTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCAAT 1119

Db        903 tgaccaccccggcaagctcatctttgctccagatcttgttctggacagggatgaggggaa 962
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1120 TGACCACCCCGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGGAA 1179

Db        963 atgcgtagaaggaattctggaaatctttgacatgctcctggcaactacttcaaggtttcg 1022
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1180 ATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTTCG 1239

Db       1023 agagttaaaactccaacacaaagaatatctctgtgtcaaggccatgatcctgctcaattc 1082
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1240 AGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAATTC 1299

Db       1083 cagtatgtaccctctggtcacagcgacccaggatgctgacagcagccggaagctggctca 1142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1300 CAGTATGTACCCTCTGGTCACAGCGACCCAGGATGCTGACAGCAGCCGGAAGCTGGCTCA 1359

Db       1143 cttgctgaacgccgtgaccgatgctttggtttgggtgattgccaagagcggcatctcctc 1202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1360 CTTGCTGAACGCCGTGACCGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATCTCCTC 1419

Db       1203 ccagcagcaatccatgcgcctggctaacctcctgatgctcctgtcccacgtcaggcatgc 1262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1420 CCAGCAGCAATCCATGCGCCTGGCTAACCTCCTGATGCTCCTGTCCCACGTCAGGCATGC 1479

Db       1263 gagtaacaagggcatggaacatctgctcaacatgaagtgcaaaaatgtggtcccagtgta 1322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1480 GAGTAACAAGGGCATGGAACATCTGCTCAACATGAAGTGCAAAAATGTGGTCCCAGTGTA 1539

Db       1323 tgacctgctgctggagatgctgaatgcccacgtgcttcgcgggtgcaagtcctccatcac 1382
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1540 TGACCTGCTGCTGGAGATGCTGAATGCCCACGTGCTTCGCGGGTGCAAGTCCTCCATCAC 1599

Db       1383 ggggtccgagtgcagcccggcagaggacagtaaaagcaaagagggctcccagaacccaca 1442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1600 GGGGTCCGAGTGCAGCCCGGCAGAGGACAGTAAAAGCAAAGAGGGCTCCCAGAACCCACA 1659

Db       1443 gtctcagtgacgcctggccctgaggtgaactggcccacagaggtcacaagctgaagcgtg 1502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1660 GTCTCAGTGACGCCTGGCCCTGAGGTGAACTGGCCCACAGAGGTCACAAGCTGAAGCGTG 1719

Db       1503 aactccagtgtgtcaggagcctgggcttcatctttctgctgtgtggtccctcatttgg 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1720 AACTCCAGTGTGTCAGGAGCCTGGGCTTCATCTTTCTGCTGTGTGGTCCCTCATTTGG 1777


RESULT      2
LOCUS       RNU57439       2555 bp    mRNA             ROD       12-JUN-1996
DEFINITION  Rattus norvegicus estrogen receptor beta mRNA, complete cds.
ACCESSION   U57439
NID         g1373280
KEYWORDS    .
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
            Murinae; Rattus.
REFERENCE   1  (bases 8 to 2555)
```

```
  AUTHORS   Kuiper,G.G., Enmark,E., Pelto-Huikko,M., Nilsson,S. and
            Gustafsson,J.A.
  TITLE     Cloning of a novel receptor expressed in rat prostate and ovary
  JOURNAL   Proc  Natl  Acad  Sci  U.S.A. 93 (12), 5925-5930 (1996)
  MEDLINE   96234066
REFERENCE   2  (bases 1 to 2555)
  AUTHORS   Enmark,E.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-MAY-1996) Eva L. Enmark, Center for Biotechnology,
            NOVUM, Karolinska, Institute, NOVUM, Huddinge, S-141 57, Sweden
FEATURES             Location/Qualifiers
     source          1..2555
                     /organism="Rattus norvegicus"
                     /strain="Sprague Dawley"
                     /sex="male"
                     /tissue_type="prostate"
     CDS             417..1874
                     /note="nuclear receptor superfamily member"
                     /codon_start=1
                     /product="estrogen receptor beta"
                     /db_xref="PID:g1373291"
                     /translation="MTFYSPAVMNYSVPGSTSNLDGGPVRLSTSPNVLWPTSGHLSPL
                     ATHCQSSLLYAEPQKSPWCEARSLEHTLPVNRETLKRKISGSSCASPVTSPNAKRDAH
                     FCPVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACR
                     LRKCYEVGMVKCGSRRERCGYRIVRRQRSSSEQVHCLSKAKRNGGHAPRVKELLLSTL
                     SPEQLVLTLLEAEPPNVLVSRPSMPFTEASMMMSLTKLADKELVHMIGWAKKIPGFVE
                     LSLLDQVRLLESCWMEVLMVGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDM
                     LLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLASANQEAESSRKLTHLLNAVTDAL
                     VWVIAKSGISSQQQSVRLANLLMLLSHVRHISNKGMEHLLSMKCKNVVPVYDLLLEML
                     NAHTLRGYKSSISGSECSSTEDSKNKESSQNLQSQ"
     misc_feature    726..923
                     /note="encodes DNA binding domain"
BASE COUNT      630 a    652 c    660 g    613 t
ORIGIN

  Query Match            60.0%;  Score 1139;  DB 92;  Length 2555;
  Best Local Similarity  84.8%;  Pred. No. 0.00e+00;
  Matches  1418;  Conservative    0;  Mismatches 250;  Indels    5;  Gaps    4;

Db     223 ataatgtccatctgtacctcttctcacaaggagttttctcagctgcgaccctctgaagac 282
           |||||| || | ||| ||||||||  ||||| ||||||||||||    | | ||  |||||
Qy      21 ATAATGACCTT-TGTGCCTCTTCTTGCAAGGTGTTTTCTCAGCTGTTATC-TC--AAGAC 76

Db     283 atggagatcaaaaactcaccgtcgagccttagttc-cctgcttcctataactgtagccag 341
           ||||| || |||||||:||| || |||||||| ||| ||| | |||||| ||||| || ||
Qy      77 ATGGATATAAAAAACTCACCATCTAGCCTTAATTCTCCTTCCTCCTACAACTGCAGTCAA 136

Db     342 tccatcctacccctggagcacggcccatctacatcccttcctcctacgtagacagcagc 401
           |||||| ||||||||||||||||| |||| ||||| ||||||||||| ||||||| || |
Qy     137 TCCATCTTACCCCTGGAGCACGGTCCATATACATACCTTCCTCCTATGTAGACAGCCAC 196

Db     402 catgagtattcagctatgacattctacagtcctgctgtgatgaactacagtgttcccggc 461
           ||||| ||| |||| |||||||||||| || |||||||||||||| |||||  ||||| ||
Qy     197 CATGAATATCCAGCCATGACATTCTATAGCCCTGCTGTGATGAATTACAGCATTCCCAGC 256

Db     462 agcaccagtaacctggacggtgggcctgtccgactgagcacaagcccaaatgtgctatgg 521
           |     || |||| |||| ||||||||||    || | || |||||||||||||||| | |||
Qy     257 AATGTCACTAACTTGGAAGGTGGGCCTGGTCGGCAGACCACAAGCCCAAATGTGTTGTGG 316

Db     522 ccaacttctgggcacctgtctcctttagcgacccattgccaatcatcgctcctctatgca 581
           |||||  |||||||||| ||||||||||| |  |||| |||| | ||| |   || |||||
Qy     317 CCAACACCTGGGCACCTTTCTCCTTTAGTGGTCCATCGCCAGTTATCACATCTGTATGCG 376

Db     582 gaacctcaaaagagtccttggtgtgaagcaagatcactagagcacacccttacctgtaaac 641
           |||||||||||||||| ||||||||||||||||||| ||||| |||||||||||||||||
Qy     377 GAACCTCAAAAGAGTCCCTGGTGTGAAGCAAGATCGCTAGAACACACCTTACCTGTAAAC 436

Db     642 agagagacactgaagaggaagcttagtgggagcagttgtgccagccctgttactagtcca 701
           ||||||||||||||| |||||| ||||||||| | |||| |||||||||||||| |||||
Qy     437 AGAGAGACACTGAAAAGGAAGGTTAGTGGGAACCGTTGCGCCAGCCCTGTTACTGGTCCA 496

Db     702 aacgcaaaggaggcatgctcacttctgcccagtctgcagcgattatgcatctggtatcat 761
               ||||||||||||||||||||| | ||||||||||||| ||||| || |||||
Qy     497 GGTTCAAAGAGGGATGCTCACTTCTGCCTGTCTGCAGCGATTACGCATCGGATATCAC 556

Db     762 tacggcgtttggtcatgtgaaggatgtaaggcctttttaaaagaagcattcaaggacat 821
           || || || ||||| ||||||||||||||||||||||||||||||||||||||||||||||
Qy     557 TATGGAGTCTGGTCGTGTGAAGGATGTAAGGCCTTTTTAAAAGAAGCATTCAAGGACAT 616

Db     822 aatgattatatctgtccagccacgaatcagtgtaccatagacaagaaccggcgtaaaagc 881
           |||||||||||| |||||||| || |||||||||| ||  || || ||||||| || |||
Qy     617 AATGATTATATTTGTCCAGCTACAAATCAGTGTACAATCGATAAAAACCGGCCCAAGAGC 676

Db     882 tgccaggcctgccgacttcgcaagtgttatgaagtaggaatggtcaagtgtggatcccgg 941
           ||||||||||||||||||||| |||||||| ||||| ||||||||| ||||||| ||  ||
Qy     677 TGCCAGGCCTGCCGACTTCGGAAGTGTTACGAAGTGGGAATGGTGAAGTGTGGCTCCCGG 736

Db     942 agagaacggtgtgggtaccgtatagtgcggaggcagagaagttctagcgagcaggtacac 1001
           |||||   | |||||||||||   | |||||||| ||||||||| |    ||||||| | |||
Qy     737 AGAGAGAGATGTGGGTACCGCCTTGTGCGGAGACAGAGAAGTGCCGACGAGCAGCTCCAC 796

Db    1002 tgcctgagcaaagccaagagaaacggtgggcatgcacccgggtgaaggagctactgctg 1061
           ||      |||| ||||||||||    || || || || ||||| |||   |||||| ||||||
Qy     797 TGTGCCGGCAAGGCCAAGAGAAGTGGCGGCCACGCGCCCCGAGTGCGGGAGCTGCTGCTG 856

Db    1062 agcaccttgagtccagagcaacttgtgctcaccctcctggaagctgaaccacccaatgtg 1121
             | || |||| || ||||| || |||||||||||||||||| ||||| |  ||| |||||
Qy     857 GACGCCCTGAGCCCCGAGCAGCTAGTGCTCACCCTCCTGGAGGCTGAGCCGCCCCATGTG 916

Db    1122 ctggtgagccgtcccagcatgcccttcaccgaggcctccatgatgatgtccctcactaag 1181
           ||| | ||||| |||||      ||||||||||||||||||||||||||||||| || |||
Qy     917 CTGATCAGCCGCCCCAGTGCGCCCTTCACCGAGGCCTCCATGATGATGTCCCTGACCAAG 976

Db    1182 ctggcggacaaggaactggtgcacatgattggctgggccaagaaaatccctggctttgtg 1241
             |||| |||||||| |||| || |||||  |||||||||||||| || || ||||| |||
Qy     977 TTGGCCGACAAGGAGTTGGTACACATGATCAGCTGGGCCAAGAAGATTCCCGGCTTTGTG 1036

Db    1242 gagctcagcctgttggaccaagtccggctcttagaaagctgctggatggaagtgctaatg 1301
           ||||||||||||| |||||||| |||||||| || |||| |||||||| || |||||||
Qy    1037 GAGCTCAGCCTGTTCGACCAAGTGCGGCTCTTGGAGAGCTGTTGGATGGAGGTGTTAATG 1096

Db    1302 gtgggactgatgtggcgctccatcgaccaccccggcaagctcatttttgctcccgacctc 1361
            |||| |||||||||||||| || |||||||||||||||||||| || ||||| || ||
Qy    1097 ATGGGGCTGATGTGGCGCTCAATTGACCACCCCGGCAAGCTCATCTTTGCTCCAGATCTT 1156

Db    1362 gttctggacagggatgaggggaagtgcgtagaagggattctggaaatctttgacatgctc 1421
           |||||||||||||||||||||| |||||||||||| |||||||||||||||||||||||||
Qy    1157 GTTCTGGACAGGGATGAGGGGAAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTC 1216

Db    1422 ctggcgacgacgtcaaggttccgtgagttaaaactccagcacaaggagtatctctgtgtg 1481
           ||||| || || |||||||| || |||||||||||||| ||||| || |||| |||||||
Qy    1217 CTGGCAACTACTTCAAGGTTTCGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTC 1276

Db    1482 aaggccatgatcctcctcaactccagtatgtaccccttggcttctgcaaaccaggaggca 1541
           |||||||||||||| ||||| ||||||||||||||   |||     | || | |||||| ||
Qy    1277 AAGGCCATGATCCTGCTCAATTCCAGTATGTACCCTCTGGTCACAGCGACCCAGGATGCT 1336

Db    1542 gaaagtagccggaagctgacacacctactgaacgcggtgacagatgccctggtctgggtg 1601
           || || ||||||||||||| ||| | |||||||| ||||| |||||   |||| ||||||
Qy    1337 GACAGCAGCCGGAAGCTGGCTCACTTGCTGAACGCCGTGACCGATGCTTTGGTTTGGGTG 1396

Db    1602 attgcgaagagtggtatctcctcccagcagcagtcagtccgactggccaacctcctgatg 1661
           ||||| ||||| || ||||||||||||||||| ||  | || ||||| |||||||||||||
Qy    1397 ATTGCCAAGAGCGGCATCTCCTCCCAGCAGCAATCCATGCGCCTGGCTAACCTCCTGATG 1456

Db    1662 cttctttctcacgtcaggcacatcagtaacaagggcatggaacatctgctcagcatgaag 1721
           || || || |||||||||||     ||||||||||||||||||||||||||| |||||||
Qy    1457 CTCCTGTCCCACGTCAGGCATGCGAGTAACAAGGGCATGGAACATCTGCTCAACATGAAG 1516

Db    1722 tgcaaaaatgtggtcccggtgtatgacctgctgctggagatgctgaatgctcacacgctt 1781
           |||||||||||||||||| ||||||||||||||||||||||||||||||||| ||   ||||
Qy    1517 TGCAAAAATGTGGTCCCAGTGTATGACCTGCTGCTGGAGATGCTGAATGCCCACGTGCTT 1576

Db    1782 cgagggtacaagtcctcaatctcggggtctgagtgcagctcaacagaggacagtaagaac 1841
           || |||| ||||||||| ||| ||||||| ||||||||| |   |||||||||||||| | |
Qy    1577 CGCGGGTGCAAGTCCTCCATCACGGGGTCCGAGTGCAGCCCGGCAGAGGACAGTAAAAGC 1636
```

```
Db   1842 aaagagggctcccagaacctacagtctcagtgatgcccagccctgaggtgaac 1894
          |||||| |||||||||||| |||||||||||| | |   | ||||||| | ||
Qy   1637 AAAGAGGGCTCCCAGAACCCACAGTCTCAGTGACGCCTGGCCCTGAGGTGAAC 1689


RESULT      3
LOCUS       CJERB          332 bp    RNA              PRI       14-NOV-1996
DEFINITION  C.jacchus mRNA for estrogen receptor beta.
ACCESSION   Y09372
NID         g1669661
KEYWORDS    DNA binding; estrogen receptor; steroid hormone receptor;
            transactivation.
SOURCE      common marmoset.
  ORGANISM  Callithrix jacchus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Platyrrhini; Callitrichidae;
            Callithrix
REFERENCE   1  (bases 1 to 332)
  AUTHORS   Gaughan,J.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 332)
  AUTHORS   Gaughan,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (11 NOV 1996) J. Gaughan, Medical Research Council,
            Centre for Reproductive Biology, 37 Chalmers Street, Edinburgh, EH3
            9EW, UK
COMMENT     Related entries: X99101 & U57439
            Related entries: X99101 & U57439.
FEATURES             Location/Qualifiers
     source          1..332
                     /organism="Callithrix jacchus"
                     /tissue_type="testis"
                     /tissue_type="ovary"
                     /tissue_type="prostate"
                     /clone="pGEM T vector"
     CDS             18..>332
                     /function="steroid hormone receptor"
                     /codon_start=1
                     /product="estrogen receptor beta"
                     /db_xref="PID:e281453"
                     /db_xref="PID:g1669662"
                     /translation="MNYSIPSSVINLEAGPGPQIISENMLWSIPGHLSPLAVHHQLSH
                     LYAEPQKSPWCEARSLEHTLPVSRETLKRKVSGNRCASPVTGPGSKRDAHFCAVCSDY
                     AAE"
BASE COUNT       91 a     84 c     77 g     80 t
ORIGIN

  Query Match             14.5%;  Score 277;  DB 70;  Length 332;
  Best Local Similarity  92.6%;  Pred. No. 3.41e-258;
  Matches  301;  Conservative    0;  Mismatches  24;  Indels    0;  Gaps    0;

Db      2 ctatagccctgctgtgatgaattacagcattcccagcagtgttactaacttggaagcggg 61
          ||||||||||||||||||||||||||||||||||||||| ||| |||||||||||||   ||
Qy    220 CTATAGCCCTGCTGTGATGAATTACAGCATTCCCAGCAATCTCACTAACTTGGAAGCTGG 279

Db     62 gcctggtcggcagatcacaagcccaaatatgctgtggtcaacacctggccacctttctcc 121
          |||||||| |||||||||||||||||||| || ||||| |||||||||| ||||||||||
Qy    280 GCCTGGTGGGCAGACCACAAGCCCAAATGTGTTGTGGCCAACACCTGGGCACCTTTCTCC 339

Db    122 tttagctgtccatcaccagttatcacatctgtatgcagaacctcaaaagactccctggtg 181
          |||||  ||||||| |||||||||||||||||||||| |||||||||||  |||||||||
Qy    340 TTTAGTGGTCCATCGCCAGTTATCACATCTGTATGCGGAACCTCAAAACAGTCCCTGGTG 399

Db    182 tgaagcaagatcactagaacataccttacctgtaagcagagagacactgaaaaggaaagt 241
          |||||||||||| |||||||| ||||||||||||| ||||||||||||||||||||| ||
Qy    400 TGAAGCAAGATCGCTAGAACACACCTTACCTGTAAACAGAGAGACACTGAAAAGGAAGGT 459

Db    242 tagtgggaaccattgtgcccagccctgttactggtccaagttcaaagagggatgctcactt 301
          ||||||||||| ||| ||||||||||||||||||||||| |||||||||||||||||||||
Qy    460 TAGTGGGAACCGTTGCGCCCAGCCCTGTTACTGGTCCAGGTTCAAAGAGGGATGCTCACTT 519

Db    302 ctgtgcagtctgcagcgattatgca 326
          ||| || |.|||||||||.||| |||
Qy    520 CTGCGCTGTCTGCAGCGATTACGCA 544


RESULT      4
LOCUS       ORZMER        1728 bp    mRNA             VRT       22-DEC-1994
DEFINITION  Killifish mRNA for estrogen receptor, complete cds.
ACCESSION   D28954
NID         g473547
KEYWORDS    estrogen receptor.
SOURCE      Oryzias sp. (strain d-rR) (library: lambda ZAP II) liver cDNA to
            mRNA.
  ORGANISM  Oryzias sp.
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Actinopterygii; Neopterygii; Teleostei; Euteleostei;
            Acanthopterygii; Atherinomorpha; Cyprinodontiformes;
            Adrianichthyoidei; Oryziidae; Oryzias.
REFERENCE   1  (bases 1 to 1728)
  AUTHORS   Okada,H.
  JOURNAL   Unpublished (1994)
COMMENT     Submitted (15-Mar-1994) to DDBJ by:
            Hisashi Okada
            Center for Gene Science
            Hiroshima University
            1-4-2 Kagamiyama, Higashi-Hiroshima
            Hiroshima 724
            Japan
            Phone: 0824-24-6283
            Fax:   0824-22-7184.
FEATURES             Location/Qualifiers
     source          1..1728
                     /organism="Oryzias sp."
                     /strain="d-rR"
                     /tissue_type="liver"
                     /clone_lib="lambda ZAP II"
     CDS             1..1728
                     /gene="MER"
                     /codon_start=1
                     /product="estrogen receptor"
                     /db_xref="PID:g602326"
                     /translation="MYPEESRGSGGVLLWIPWKGRTTMPAPNPATTPLYSQSSTGYYS
                     APLEINGPPSEGSLQSLGSGPISPLVFVPSSPRLSPFMHPPSHHYLETTSTPVYRSSH
                     QGASREDQCGSREDTCSLGELGAGAGAGGFEMAKDTRFCAVCSDYASGYHYGVWSCEG
                     CKAFFKRSIQGHNDYMCPATNQCTIDRNRRKSCQACRLRKCYEVGMMKGGVRKDRIRI
                     LRPDKRRTGVGDGDKVVKGQEHKIVHYDGRKRSSIGGGGGGGGRLSVTSIPPPQVLL
                     LLQGAEPPILCSRQKLSRPYTEVTMMTLLTSMADKELVHMIAWAKKIPGFLQLSLHDQ
                     VLLLESSWLEVLMIGLIWRSIHCPGKLIFAQDLILDRNEGDCVEGMTEIFDMLLATAS
                     RFRVLKLKPEEFVCLKAIILLNSGAFSFCTGIMEPLHNSAAVQSMLDTITDALIHYIS
                     QSGYLAQEQARRQAQLLLLLSHIRHMSNKGMEHLYSMKCKNKVPLYDLLLEMLDAHRL
                     HHPVRAPQSLSQVDRDPPSTSSGGGGIAPGSISASRGRIESPSRGPFAPSVLQYGGSR
                     PDCIPALQD"
BASE COUNT      379 a    513 c    496 g    340 t
ORIGIN

  Query Match             11.4%;  Score 217;  DB 51;  Length 1728;
  Best Local Similarity  68.2%;  Pred. No. 8.99e-193;
  Matches  434;  Conservative    0;  Mismatches 199;  Indels    3;  Gaps    1;

Db    883 ccgtacaccgaggtcaccatgatgaccctgctcaccagcatggcagacaaggagctggtc 942
          || | |||||||| | |||||||||     || ||||    |||| |||||||||| ||||
Qy    938 CCCTTCACCGAGGCTCCATCATCATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTA 997

Db    943 cacatgatcgcctgggccaagaagctcccaggttttctgcagctgtccctgcacgatcag 1002
          |||||||||  ||||||||||||| | || ||| || ||||    ||||   ||| ||
Qy    998 CACATGATCAGCTGGGCCAAGAAGATTCCCGGCTTTCTGGAGCTCAGCCTGTTCGACCAA 1057

Db   1003 gtgctgctgctggagagctcgtggctggaggtgctcatgatcgggctcatttggaggtcc 1062
          |||| |||  |||||||||  ||| |||||||| | ||||| || || || ||| | ||
Qy   1058 GTGCGGCTGCTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCA 1117

Db   1063 atccactgtcccgggaagctcatctttgcacaagacctcatcctggacaggaatgagggg 1122
          ||  ||   ||||| |||||||||||||| | ||| ||   | ||||||||| |||||||
Qy   1118 ATTGACCACCCCGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGG 1177
```

```
Db    1123 gactgcgtggaaggcatgacggagatcttcgacatgctgctggcaactgcttcccgcttc 1182
            | ||||| ||||| ||   ||| ||||| |||||||| ||||||||| ||||  | ||
Qy    1178 AAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTT 1237

Db    1183 cgtgtgctcaaactcaaacctgaggaattcgtctgcctcaaagctattattttactcaac 1242
           || | | | |||||| |||   | ||||   ||||  |||| || || ||  | |||||
Qy    1238 CGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATCATCCTGCTCAAT 1297

Db    1243 tcggtgctttttctttctgcaccggcaccatggagccacttcacaacagcgcggcggtt 1302
           ||| ||   |   || |   ||| |  |||  |||  |   | ||| ||||  |  | |
Qy    1298 TCCAGTATGTACCCTCTGGTCACAGCGACCCAGGATGC---TGATAGCAGCCGGAAGCTG 1354

Db    1303 cagagcatgctggacaccatcacagacgcactcattcattacatcagtcagtcgggttac 1362
                | ||||| || || | || || ||  |  ||      ||     ||   ||   |
Qy    1355 GCTCACTTGCTGAACGCCGTGACCGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATC 1414

Db    1363 ttggcccaggagcaggcgagacggcaggcccagctgctcctgctgctctcccacatcagg 1422
               ||||| ||||  | |  || | |||  | || ||  |||| || |||||| |||||
Qy    1415 TCCTCCCAGCAGCAATCCATGCGCCTGGCTAACCTCCTGATGCTCCTGTCCCACGTCAGG 1474

Db    1423 cacatgagcaacaaaggcatggagcacctctacagcatgaagtgcaagaacaaagtccct 1482
           ||   ||| ||||| |||||||| || ||   ||  ||||||||||| ||    |||||
Qy    1475 CATGCGAGTAACAAGGGCATGGAACATCTGCTCAACATGAAGTGCAAAAATGTGGTCCCA 1534

Db    1483 ctttatgacctcctactggagatgctcgatgcccac 1518
            | |||||||| || |||||||||||  ||||||||
Qy    1535 GTGTATGACCTGCTGCTGGAGATGCTGAATGCCCAC 1570


RESULT      5
LOCUS       SSESTREC     1788 bp    RNA             MAM       15-SEP-1994
DEFINITION  S.scrofa mRNA for estradiol receptor.
ACCESSION   Z37167
NID         g587554
KEYWORDS    estradiol receptor
SOURCE      pig.
  ORGANISM  Sus scrofa
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Artiodactyla; Suiformes; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 1788)
  AUTHORS   Bokenkamp,D., Jungblut,P.W. and Thole,H.H.
  TITLE     The C-terminal half of the porcine estradiol receptor contains no
            post-translational modification: determination of the primary
            structure
  JOURNAL   Mol. Cell. Endocrinol. 104 (2), 165-172 (1994)
  MEDLINE   95080454
REFERENCE   2  (bases 1 to 1788)
  AUTHORS   Boekenkamp,D.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-SEP-1994) Dirk Boekenkamp, Max-Planck-Institut fuer
            experimentelle Endokrinologie, Feodor-Lynen Strasse 7, Hannover,
            30625, Germany
FEATURES             Location/Qualifiers
     source          1..1788
                     /organism="Sus scrofa"
                     /tissue_type="Uterus"
                     /sex="Female"
     CDS             1..1788
                     /codon_start=1
                     /product="estradiol receptor"
                     /db_xref="PID:g587555"
                     /translation="MTMTLHTKASGMALLHQIQANELEPLNRPQLKIPLERPLGEVYV
                     DSSKPAVYNYPEGAAYDFNAAAAASAPVYGQSGLAYGPGSEAAAFGANGLGGFQPLNS
                     VSPSPLVLLHPPPQLSPFLHPHGQQVPYYIENEPSGYAVREAGPPAFYRPNSDNRRQG
                     GRERLASTSDKGSMAMESAKETRYCAVCNDYASGYHYGVWSCEGCKAFFKRSIQGHND
                     YMCPATNQCTIDKNRRKSCQACRLRKCYEVGMMKGGIRKDRRGGRMLKHKRQRDDGEG
                     RNEAVPPGDMRSANLWPSPLLIKHTKKNSPVLSLTADQMISALLEAEPPIIYSEYDPT
                     RPLSEASMMGLLTNLADRELVHMINWAKRVPGFLDLSLHDQVHLLECAWLEILMIGLV
                     WRSMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEFVCLKS
                     IILLNSGVYTFLSSTLKSLEEKDHIHRVLDKITDTLIHLMAKAGLTLQQLHRRLAQLL
                     LILSHFRHMSNKGMEHLYNMKCKNVVPLYDLLLEMLDAHRLHAPTNLGGPPPEDMSQS
                     QLATSGSTPSHSLQMYYITGEAENFPTTI"
BASE COUNT      410 a    530 c    504 g    344 t
ORIGIN

  Query Match               10.7%;  Score 203;  DB 48;  Length 1788;
  Best Local Similarity     67.1%;  Pred. No. 1.24e-177;
  Matches  427;  Conservative    0;  Mismatches  206;  Indels    3;  Gaps    2;

Db    1006 gccctcagtgaggcttcaatgatgggcttgctgaccaacctcgcagacagggagctggta 1065
           ||| |||  ||||| || ||||||   |  ||||||||  | || |||| |||| |||||
Qy     938 GCCTTCACCGAGGCCTCCATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTA 997

Db    1066 cacatgatcaactgggcaaagagggtgccaggattttggatttaagcctccatgatcaa 1125
           |||||||||| |||||| |||| | | || || ||| ||||  | |||||    || |||
Qy     998 CACATGATCAGCTGGGCCAAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAA 1057

Db    1126 gtgcatcttctggaatgtgcctggctagagatcctcatgattggtcttgtctggcgctcc 1185
           ||||  ||  ||||  |    ||| | ||| |  | ||||| || ||  | ||||||||
Qy    1058 GTGCGGCTCTTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCA 1117

Db    1186 atggagcacccagggaagctcctgtttgctcctaacttgctcctggacaggaaccaggga 1245
           || || ||||| || |||||| | ||||||||  |  |  | ||||||||| |  ||||
Qy    1118 ATTGACCACCCGGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGG 1177

Db    1246 aagtgtgtcgagggaatggtggagatctttgacatgttgctggctacatcatctcgcttc 1305
           || || || || |||||  |||| |||||||||||| | ||||| ||  | ||  | ||
Qy    1178 AAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTT 1237

Db    1306 cgtatgatgaatctccagggagaggagtttgtgtgcctcaaatccatcattttgcttaat 1365
           ||   | | || |||||     | || | | | ||  ||||  |||||||  |||| |||
Qy    1238 CGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATCATCCTGCTCAAT 1297

Db    1366 tctggagtgtacacgtttctgtccagcaccctgaagtctctggaagagaaggaccatatc 1425
           ||  |  ||||| |  |  |  |    |||| | |  ||    |  ||  |||   |  |
Qy    1298 TCCAGTATGTACCCTCTGGTCACAGCGACCCAGGATGCTGAT-AGCAGCCGGAAGCTGGC 1356

Db    1426 caccgtgtcctggacaagatcacagacactttgatccacctgatggccaaggcaggcctg 1485
              | ||  ||| ||    | || ||  ||||| |     |||| ||||||   ||| |
Qy    1357 TCACTTG--CTGAACGCCGTGACCGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATC 1414

Db    1486 actctgcagcagcagcaccagcgtctggcgcagctcctcctcatcctgtctcacatcagg 1545
            |    ||||||||   |  ||| |||||  | |||||  |  ||||||| ||| |||||
Qy    1415 TCCTCCCAGCAGCAATCCATGCGCCTGGCTAACCTCCTGATGCTCCTGTCCCACGTCAGG 1474

Db    1546 cacatgagtaacaaaggcatggagcatctgtacagcatgaagtgcaagaacgtggtgccc 1605
           ||   ||||||||| |||||||| ||||||  || |||||||||||| || ||||| ||
Qy    1475 CATGCGAGTAACAAGGGCATGGAACATCTGCTCAACATGAAGTGCAAAAATGTGGTCCCA 1534

Db    1606 ctctatgacctgctgctggagatgctggacgcccac 1641
            | |||||||||||||||||||||||| | ||||||
Qy    1535 GTGTATGACCTGCTGCTGGAGATGCTGAATGCCCAC 1570


RESULT      6
LOCUS       OAESTREC     1966 bp    RNA             MAM       [illegible]-1996
DEFINITION  O.aries mRNA for estrogen receptor.
ACCESSION   Z49257
NID         g1617201
KEYWORDS    estrogen receptor.
SOURCE      sheep.
  ORGANISM  Ovis aries
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Artiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Caprinae; Ovis.
REFERENCE   1  (bases 1 to 1966)
  AUTHORS   Madigou,T., Tiffoche,C., Lazennec,G., Pelletier,J. and
            Thieulant,M.L.
  TITLE     The sheep estrogen receptor: cloning and regulation of expression
            in the hypothalamo-pituitary axis
  JOURNAL   Mol. Cell. Endocrinol. 121 (2), 153-163 (1996)
  MEDLINE   97047396
REFERENCE   2  (bases 1 to 1966)
  AUTHORS   Madigou,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (12 MAY 1995) Thierry Madigou, Interactions Cellulaires
            et Moleculaires, CNRS URA 256, Rennes, 35042, FRANCE
FEATURES             Location/Qualifiers
     source          1..1966
                     /organism="Ovis aries"
                     /strain="Romanov"
                     /clone="lambda P112"
                     /dev_stage="adult"
                     /tissue_type="anterior pituitary gland"
                     /clone_lib="cDNA library"
                     /sex="Female"
     mRNA            1..1966
                     /product="estrogen receptor"
BASE COUNT      439 a    604 c    549 g    374 t
ORIGIN

  Query Match            10.6%;  Score 201;  DB 47;  Length 1966;
  Best Local Similarity  67.0%;  Pred. No. 1.78e-175;
  Matches  426;  Conservative    0;  Mismatches 207;  Indels    3;  Gaps    3;

Db    1158 cctttcagtgaggcttcaatgatgggcttgctgaccggccttgcagacagggagctggta 1217
           || ||||  ||||| || ||||||   |  ||||||    | || |||| |||| |||||
Qy     938 CCCTTCACCGAGGCCTCCATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTA 997

Db    1218 cacatgatcaactgggcaaagagagtgccaggatttgtggatctggccctccatgatcaa 1277
           |||||||||| |||||| ||||   | || || |||||||| ||   |||    || |||
Qy     998 CACATGATCAGCTGGGCCAAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAA 1057

Db    1278 gtccaccttttggaatgtgcctggctagagatcctcatgattggtcttgtctggcgctcc 1337
           || |  || |||||  |    ||| | ||| |  | ||||| || ||  | ||||||||
Qy    1058 GTGCGGCTCTTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCA 1117

Db    1338 atggagcatccagggaagctcctatttgctcccaaccttctcctggacaggaaccaggga 1397
           || || || || || |||||| | ||||||||  | ||| | ||||||||| |  ||||
Qy    1118 ATTGACCACCCCGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGG 1177

Db    1398 aaatgtgtggaaggcatggtggagatctttgacatgttgctggctacatcgtctcggttc 1457
           ||||| || ||||| ||  |||| |||||||||||| | ||||| ||  | ||  | ||
Qy    1178 AAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGCTTT 1237

Db    1458 cgtatgatgaatctccagggagaagagtttgtgtgcctcaaatccatcatcttgcttaat 1517
           ||   | | || |||||     |||| | | | ||  ||||  |||| ||| |||| |||
Qy    1238 CGAGAGCTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAAT 1297

Db    1518 tctggggtgtacacatttctgtccagcactctgaggtctctggaagagaaggaccacatc 1577
           ||  |  ||||| |  |  |  | |||   |  |||   |||  ||     |    |
Qy    1298 TCCAGTATGTACCCTCTGGTCAC-AGCGACCC-AGGATGCTGACAGCAGCCGGAAGCTGG 1355

Db    1578 caccgcgtcctggacaagatcacagacaccttgatccatctgatggccaaagcaggcctg 1637
           | |  | | ||| ||    | || ||  | ||| |     |||| |||||    ||| |
Qy    1356 CTCA-CTTGCTGAACGCCGTGACCGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATC 1414

Db    1638 actctgcagcagcagcaccggcgtctggcccagttcctcctcctcctctctcacttcagg 1697
            |    ||||||||   |  ||| |||||  |  ||||  | ||||| || ||| |||||
Qy    1415 TCCTCCCAGCAGCAATCCATGCGCCTGGCTAACCTCCTGATGCTCCTGTCCCACGTCAGG 1474

Db    1698 cacatgagcaacaaaggcatggagcacctatacagcatgaagtgcaagaacgtggtgccc 1757
           ||   ||| ||||| |||||||| || ||   || |||||||||||| || ||||| ||
Qy    1475 CATGCGAGTAACAAGGGCATGGAACATCTGCTCAACATGAAGTGCAAAAATGTGGTCCCA 1534

Db    1758 ctctacgacctgctgctggagatgctggatgcccac 1793
            | || ||||||||||||||||||||| ||||||||
Qy    1535 GTGTATGACCTGCTGCTGGAGATGCTGAATGCCCAC 1570


RESULT      7
LOCUS       GGERR         2085 bp    RNA             VRT       12-SEP-1993
DEFINITION  Chicken mRNA for estrogen receptor.
ACCESSION   X03805
NID         g63378
KEYWORDS    DNA binding protein; estrogen receptor; receptor; steroid hormone
            receptor; unidentified reading frame.
SOURCE      chicken.
  ORGANISM  Gallus gallus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Archosauria; Aves; Neognathae; Galliformes;
            Phasianidae; Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 2085)
  AUTHORS   Krust,A., Green,S., Argos,P., Kumar,V., Walter,P., Bornert,J.M. and
            Chambon,P.
  TITLE     The chicken oestrogen receptor sequence: homology with v-erbA and
            the human oestrogen and glucocorticoid receptors
  JOURNAL   EMBO J. 5 (5), 891-897 (1986)
  MEDLINE   86247578
COMMENT     Data kindly reviewed (23-JUL-1986) by S. Green.
FEATURES             Location/Qualifiers
     source          1..2085
                     /organism="Gallus gallus"
     mRNA            1..2085
                     /note="ER mRNA"
     CDS             69..131
                     /note="open reading frame 1 (aa 1-20)"
                     /codon_start=1
                     /db_xref="PID:g63379"
                     /translation="MYCCFNMISAEPQNRFWCFF"
     CDS             177..200
                     /note="open reading frame 2 (aa 1-7)"
                     /codon_start=1
                     /db_xref="PID:g584490"
                     /translation="MFLAHCH"
     CDS             203..1972
                     /note="chicken estrogen receptor (aa 1-589)"
                     /codon_start=1
                     /db_xref="PID:g63380"
                     /db_xref="SWISS-PROT:P06212"
                     /translation="MTMTLHTKASGVTLLHQIQGTELETLSRPQLKIPLERSLSDMYV
                     ESNKTGVFNYPEGATYDFGTTAPVYGSTTLSYAPTSESFGSSSLAGFHSLNNVPPSPV
                     VFLQTAPQLSPFIHHHSQQVPYYLENEQGSFGMREAAPPAFYRPSSDNRRHSIRERMS
                     STNEKGSLSMESTKETRYCAVCNDYASGYHYGVWSCEGCKAFFKRSIQGHNDYMCPAT
                     NQCTIDKNRRKSCQACRLRKCYEVGMMKGGIRKDRRGGRMMKQKRQREEQDSRNGEAS
                     STELRAPTLWTSPLVVKHNKKNSPALSLTAEQMVSALLEAEPPIVYSEYDPNRPFNEA
                     SMMTLLTNLADRELVHMINWAKRVPGFVDLTLHDQVHLLECAWLEILMIGLVWRSMEH
                     PGKLLFAPNLLLDRNQGKCVEGMVEIFDMLLATAARFRMMNLQGEEFVCLKSIILLNS
                     GVYTFLSSTLKSLEERDYIHRVLDKITDTLIHLMAKSGLSLQQQHRRLAQLLLILSHI
                     RHMSNKGMEHLYNMKCKNVVPLYDLLLEMLDAHRLHAPAARSAAPMEEENRNQLTTAP
                     ASSHSLQSFYINSKEEESMQNTI"
     misc_feature    203..313
                     /note="aa 1-37 --> highly conserved region A"
     misc_feature    314..721
                     /note="conserved region B (aa 38-173)"
     misc_feature    722..970
                     /note="highly conserved region C (aa 174-256) pot.
                     DNA-binding domain"
     misc_feature    971..1087
                     /note="conserved region D (aa 257-295)"
     misc_feature    1088..1840
                     /note="highly conserved region E (aa 296-546) certain
                     estrogen binding domain"
     misc_feature    1841..1969
                     /note="conserved region F (aa 547-589)"
BASE COUNT      582 a    519 c    480 g    504 t
ORIGIN

  Query Match            10.6%;  Score 201;  DB 50;  Length 2085;
  Best Local Similarity  66.7%;  Pred. No. 1.78e-175;
  Matches  431;  Conservative    0;  Mismatches 212;  Indels    3;  Gaps    2;

Db    1180 ccccaatagaccattcaacgaagcatctatgatgaccctgttgaccaaccttgcagacag 1239
           ||||| |   || |||| ||| || || |||||||      |||||||  | || ||||
Qy     928 CCCCAGTCCGCCCTTCACCGAGGCCTCCATGATGATGTCCCTGACCAAGTTGGCCGACAA 987

Db    1240 agaattagtgcacatgatcaactgggcaaagagagttccaggatttgtggatttaacact 1299
            || || || |||||||||| |||||| ||||   |||| || ||||||||  | |  ||
Qy     988 GGAGTTGGTACACATGATCAGCTGGGCCAAGAAGATTCCCGGCTTTGTGGAGCTCAGCCT 1047

Db    1300 ccatgatcaggtccatctgctggaatgtgcctggttagagatattgatgatcggcttagt 1359
```

```
               || ||  || |   ||  ||||   |      ||| | ||  | ||  ||||| ||   |   |
Qy   1048 GTTCGACCAAGTGCGGCTCTTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGAT 1107

Db   1360 ctggcgctccatggaacacccaggaaagcttttatttgcacctaatctattactggacag 1419
            |||||||| || || ||||| || |||||   | ||||| ||   ||||   | |||||||
Qy   1108 GTGGCGCTCAATTGACCACCCCGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAG 1167

Db   1420 gaatcaagggaaatgtgtagaaggcatggtagaaatctttgacatgctactggctactgc 1479
           | || | |||||||| ||||| || ||  |||||||||||||||||| ||||| ||| |
Qy   1168 GGATGAGGGGAAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTAC 1227

Db   1480 tactcggtttcgaatgatgaaccttcaaggggaggaatttgtatgccttaagtccatcat 1539
           | |  |||||||   | | || || |||     | |||| | | ||   | ||| |||| ||
Qy   1228 TTCAAGGTTTCGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGAT 1287

Db   1540 cctgctcaattctggtgtgtacacttttctttctagcaccttgaaatctctggaagagag 1599
          |||||||||||   || ||||| || |   |   |      |||  | |  ||    |  ||
Qy   1288 CCTGCTCAATTCCAGTATGTACCCTCTGGTCACAGCGACCCAGGATGCTGAC-AGCAGCC 1346

Db   1600 ggactatatccaccgtgttctggacaaaatcacagatactctgatacacctaatggcaaa 1659
          |||    |  |    | ||  ||| ||     | || ||| || || |       | || || ||
Qy   1347 GGAAGCTGGCTCACTTG--CTGAACGCCGTGACCGATGCTTTGGTTTGGGTGATTGCCAA 1404

Db   1660 gtcaggtctttctctgcagcagcaacaccggcgactagctcagctcctccttatcctctc 1719
          |  ||  | ||    |||||||||  |  ||| || ||| | |||||   |  |||| ||
Qy   1405 GAGCGGCATCTCCTCCCAGCAGCAATCCATGCGCCTGGCTAACCTCCTGATGCTCCTGTC 1464

Db   1720 tcacattaggcatatgagcaacaaaggaatggagcacctgtacaatatgaagtgtaaaaa 1779
            ||| ||||||||  ||| ||||| || ||||| || |||   ||| |||||||| |||||
Qy   1465 CCACGTCAGGCATGCGAGTAACAAGGGCATGGAACATCTGCTCAACATGAAGTGCAAAAA 1524

Db   1780 tgtagttccgctctacgacctcttactggagatgctggacgctcac 1825
          ||| ||  ||   | || |||||   | |||||||||||| | ||  |||
Qy   1525 TGTGGTCCCAGTGTATGACCTGCTGCTGGAGATGCTGAATGCCCAC 1570


RESULT    8
LOCUS       I08538       2092 bp                    PAT       14-NOV-1994
DEFINITION  Sequence 1 from Patent WO 8705049
ACCESSION   I08538
NID         g588753
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2092)
  AUTHORS   Shine,J.
  TITLE     EUCARYOTIC EXPRESSION OF STEROID RECEPTOR PROTEINS
  JOURNAL   Patent: WO 8705049-A 1 27-AUG-1987;

FEATURES             Location/Qualifiers
     source          1..2092
                     /organism="unknown"
BASE COUNT      473 a    605 c    593 g    421 t
ORIGIN

  Query Match            10.4%;  Score 197;  DB 55;  Length 2092
  Best Local Similarity  66.7%;  Pred. No. 3.65e-171;
  Matches  424;  Conservative    0;  Mismatches 209;  Indels    3;  Gaps    2;

Db   1298 cccttcagtgaagcttcgatgatgggcttactgaccaacctggcagacagggagctggtt 1357
          ||||||||  || || || |||||||    |  ||||||||||   |||| |||| |||| ||||
Qy    938 CCCTTCAGCGAGGCCTCCATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTA 997

Db   1358 cacatgatcaactgggcgaagagggtgccaggctttgtggatttgaccctccatgatcag 1417
          |||||||||   |||| |||| | |  || ||||||||||||    | | |||     || ||
Qy    998 CACATGATCAGCTGGGCCAAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAA 1057

Db   1418 gtccaccttctagaatgtgcctggctagagatcctgatgattggtctcgtctggcgctcc 1477
          || |  ||   | ||  |      ||| | ||| |  | ||||| || || | ||||||||
Qy   1058 GTGCGGCTCTTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCA 1117

Db   1478 atggagcacccagtgaagctactgtttgctcctaacttgctcttggacaggaaccaggga 1537
          || || ||||| |  ||||| | |||||||||    |   |   ||||||| |  |||| 
Qy   1118 ATTGACCACCCCGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGG 1177

Db   1538 aaatgtgtagagggcatggtggagatcttcgacatgctgctggctacatcatctcggttc 1597
          ||||| ||||| || ||  |||| ||||| |||||||| ||||| ||  | ||  ||||
Qy   1178 AAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTT 1237

Db   1598 cgcatgatgaatctgcagggagaggagtttgtgtgcctcaaatctattattttgcttaat 1657
          ||    | | || || ||     | || | | | || |||| | || || |||| |||
Qy   1238 CGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAAT 1297

Db   1658 tctggagtgtacacatttctgtccagcaccctgaagtctctggaagagaaggaccatatc 1717
          || | ||||| | | | |     |||| | | ||    | ||  |||   | |
Qy   1298 TCCAGTATGTACCCTCTGGTCACAGCGACCCAGGATGCTGAC-AGCAGCCGGAAGCTGGC 1356

Db   1718 caccgagtcctggacaagatcacagacactttgatccacctgatggccaaggcaggcctg 1777
              |  |  ||| ||    | || ||  ||||| |     |||| ||||||    ||| |
Qy   1357 TCACTTG--CTGAACGCCGTGACCGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATC 1414

Db   1778 accctgcagcagcagcaccagcggctggcccagctcctcctcatcctctcccacatcagg 1837
           ||  |||||||| | ||| ||||| | ||||| |   |||| |||||| |||||
Qy   1415 TCCTCCCAGCAGCAATCCATGCGCCTGGCTAACCTCCTGATGCTCCTGTCCCACGTCAGG 1474

Db   1838 cacatgagtaacaaaggcatggagcatctgtacagcatgaagtgcaagaacgtggtgccc 1897
          ||   |||||||||| ||||||||| ||||||  || ||||||||||||| || ||||  |
Qy   1475 CATGCGAGTAACAAGGGCATGGAACATCTGCTCAACATGAAGTGCAAAAATGTGGTCCCA 1534

Db   1898 ctctatgacctgctgctggagatgctggacgcccac 1933
           | |||||||||||||||||||||||| | ||||||
Qy   1535 GTGTATGACCTGCTGCTGGAGATGCTGAATGCCCAC 1570


RESULT    9
LOCUS       HUMERMCF     2092 bp    mRNA             PRI       08-NOV-1994
DEFINITION  Human estrogen receptor mRNA, complete cds.
ACCESSION   M12674
NID         g182192
KEYWORDS    estrogen receptor.
SOURCE      Human breast cancer cell line MCF-7, cDNA to mRNA, clone OR8.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2092)
  AUTHORS   Greene,G.L., Gilna,P., Waterfield,M., Baker,A., Hort,Y. and
            Shine,J.
  TITLE     Sequence and expression of human estrogen receptor complementary
            DNA
  JOURNAL   Science 231 (4742), 1150-1154 (1986)
  MEDLINE   86122927
FEATURES             Location/Qualifiers
     source          1..2092
                     /organism="Homo sapiens"
                     /cell_line="MCF-7"
                     /cell_type="Mammary Carcinoma"
                     /tissue_lib="lambda gt11"
                     /lab_host="Prokaryotae"
                     /map="6q24-q27"
     mRNA            <1..>2092
                     /note="ER mRNA"
     CDS             293..2080
                     /gene="ESR"
                     /note="estrogen receptor"
                     /codon_start=1
                     /db_xref="GDB:G00-119-120"
                     /db_xref="PID:g182193"
                     /translation="MTMTLHTKASGMALLHQIQGNELEPLNRPQLKIPLERPLGEVYL
                     DSSKPAVYNYPEGAAYEFNAAAAANAQVYGQTGLPYGPGSEAAAFGSNGLGGFPPLNS
                     VSPSPLMLLHPPPQLSPFLQPHGQQVPYYLENEPSGYTVREAGPPAFYRPNSDNRRQG
                     GRERLASTNDKGSMAMESAKETRYCAVCNDYASGYHYGVWSCEGCKAFFKRSIQGHND
                     YMCPATNQCTIDKNRRKSCQACRLRKCYEVGMMKGGIRKDRRGGRMLKHKRQRDDGEG
                     RGEVGSAGDMRAANLWPSPLMIKRSKKNSLALSLTADQMVSALLDAEPPILYSEYDPT
                     RPFSEASMMGLLTNLADRELVHMINWAKRVPGFVDLTLHDQVHLLECAWLEILMIGLV
                     WRSMEHPVKLLFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEFVCLKS
                     IILLNSGVYTFLSSTLKSLEEKDHIHRVLDKITDTLIHLMAKAGLTLQQQHQRLAQLL
                     LILSHIRHMSNKGMEHLYSMKCKNVVPLYDLLLEMLDAHRLHAPTSRGGASVEETDQS
                     HLATAGSTSSHSLQKYYITGEAEGFPATV"
BASE COUNT      473 a    605 c    593 g    421 t
ORIGIN      1 bp upstream of EcoRI site.

  Query Match             10.4%;  Score 197;  DB 80;  Length 2092;
  Best Local Similarity   66.7%;  Pred. No. 3.65e-171;
  Matches  424;  Conservative    0;  Mismatches 209;  Indels    3;  Gaps    2;

Db   1298 cccttcagtgaagcttcgatgatgggcttactgaccaacctggcagacagggagctggtt 1357
          |||||||  || || || ||||||   |  ||||||||  |||| |||| |||| ||||
Qy    938 CCCTTCACCGAGGCCTCCATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTA 997

Db   1358 cacatgatcaactgggcgaagagggtgccaggctttgtggatttgaccctccatgatcag 1417
          |||||||||| |||||| |||| | | || |||||||||||  | | |||    || ||
Qy    998 CACATGATCAGCTGGGCCAAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAA 1057

Db   1418 gtccaccttctagaatgtgcctggctagagatcctgatgattggtctcgtctggcgctcc 1477
          || |  ||  | ||  |    ||| | ||| |  | ||||| || ||  | ||||||||
Qy   1058 GTGCGGCTCTTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCA 1117

Db   1478 atggagcacccagtgaagctactgtttgctcctaacttgctcttggacaggaaccaggga 1537
          || || ||||| |  |||||  | ||||||||  |  |  |  |||||||| |  ||||
Qy   1118 ATTGACCACCCCGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGG 1177

Db   1538 aaatgtgtagagggcatggtggagatcttcgacatgctgctggctacatcatctcggttc 1597
          ||||| ||||| || ||  |||| ||||| |||||||| ||||| ||  | ||  ||||
Qy   1178 AAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTT 1237

Db   1598 cgcatgatgaatctgcagggagaggagtttgtgtgcctcaaatctattattttgcttaat 1657
          ||   | | || || ||     | || | | | ||  ||||  | || ||  |||| |||
Qy   1238 CGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAAT 1297

Db   1658 tctggagtgtacacatttctgtccagcaccctgaagtctctggaagagaaggaccatatc 1717
          ||  |  ||||| |  |  |  |    |||| | |  ||    |  ||  |||   |  |
Qy   1298 TCCAGTATGTACCCTCTGGTCACAGCGACCCAGGATGCTGAC-AGCAGCCGGAAGCTGGC 1356

Db   1718 caccgagtcctggacaagatcacagacactttgatccacctgatggccaaggcaggcctg 1777
             |  |  ||| ||    | || ||  ||||| |     |||| ||||||   ||| |
Qy   1357 TCACTTG--CTGAACGCCGTGACCGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATC 1414

Db   1778 accctgcagcagcagcaccagcggctggcccagctcctcctcatcctctcccacatcagg 1837
           ||   ||||||||   |  ||| |||||  | |||||  |  |||| |||||| |||||
Qy   1415 TCCTCCCAGCAGCAATCCATGCGCCTGGCTAACCTCCTGATGCTCCTGTCCCACGTCAGG 1474

Db   1838 cacatgagtaacaaaggcatggagcatctgtacagcatgaagtgcaagaacgtggtgccc 1897
          ||   ||||||||| |||||||| ||||||  || |||||||||||| || ||||| ||
Qy   1475 CATGCGAGTAACAAGGGCATGGAACATCTGCTCAACATGAAGTGCAAAAATGTGGTCCCA 1534

Db   1898 ctctatgacctgctgctggagatgctggacgcccac 1933
           | |||||||||||||||||||||||| | ||||||
Qy   1535 GTGTATGACCTGCTGCTGGAGATGCTGAATGCCCAC 1570


RESULT    10
LOCUS       I15368          4963 bp    DNA             PAT       09-NOV-1995
DEFINITION  Sequence 16 from patent US 5464758.
ACCESSION   I15368
NID         g1250276
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 4963)
  AUTHORS   Gossen,M. and Bujard,H.
  TITLE     Tight control of gene expression in eucaryotic cells by
            tetracycline-responsive promoters
  JOURNAL   Patent: US 5464758-A 16 07-NOV-1995;

FEATURES             Location/Qualifiers
     source          1..4963
                     /organism="unknown"
```

```
BASE COUNT     1259 a   1300 c   1278 g   1126 t
ORIGIN

  Query Match             10.4%;  Score 197;  DB 56;  Length 4963;
  Best Local Similarity   66.7%;  Pred. No. 3.65e-171;
  Matches  424;  Conservative    0;  Mismatches 209;  Indels    3;  Gaps    2;

Db   1471 cccttcagtgaagcttcgatgatgggcttactgaccaacctggcagacagggagctggtt 1530
          |||||||  || || || ||||||   |  ||||||||  |||| |||| |||| ||||
Qy    938 CCCTTCACCGAGGCCTCCATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTA 997

Db   1531 cacatgatcaactgggcgaagagggtgccaggctttgtggatttgaccctccatgatcag 1590
          |||||||||| |||||| |||| | | || |||||||||||  | | |||    || ||
Qy    998 CACATGATCAGCTGGGCCAAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAA 1057

Db   1591 gtccaccttctagaatgtgcctggctagagatcctgatgattggtctcgtctggcgctcc 1650
          || |  ||  | ||  |    ||| | ||| |  | ||||| || ||  | ||||||||
Qy   1058 GTGCGGCTCTTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCA 1117

Db   1651 atggagcacccagtgaagctactgtttgctcctaacttgctcttggacaggaaccaggga 1710
          || || ||||| |  |||||  | ||||||||  |  |  |  |||||||| |  ||||
Qy   1118 ATTGACCACCCCGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGG 1177

Db   1711 aaatgtgtagagggcatggtggagatcttcgacatgctgctggctacatcatctcggttc 1770
          ||||| ||||| || ||  |||| ||||| |||||||| ||||| ||  | ||  ||||
Qy   1178 AAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTT 1237

Db   1771 cgcatgatgaatctgcagggagaggagtttgtgtgcctcaaatctattattttgcttaat 1830
          ||   | | || || ||     | || | | | ||  ||||  | || ||  |||| |||
Qy   1238 CGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAAT 1297

Db   1831 tctggagtgtacacatttctgtccagcaccctgaagtctctggaagagaaggaccatatc 1890
          ||  |  ||||| |  |  |  |    |||| | |  ||    |  ||  |||   |  |
Qy   1298 TCCAGTATGTACCCTCTGGTCACAGCGACCCAGGATGCTGAC-AGCAGCCGGAAGCTGGC 1356

Db   1891 caccgagtcctggacaagatcacagacactttgatccacctgatggccaaggcaggcctg 1950
             |  |  ||| ||    | || ||  ||||| |     |||| ||||||   ||| |
Qy   1357 TCACTTG--CTGAACGCCGTGACCGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATC 1414

Db   1951 accctgcagcagcagcaccagcggctggcccagctcctcctcatcctctcccacatcagg 2010
           ||   ||||||||   |  ||| |||||  | |||||  |  |||| |||||| |||||
Qy   1415 TCCTCCCAGCAGCAATCCATGCGCCTGGCTAACCTCCTGATGCTCCTGTCCCACGTCAGG 1474

Db   2011 cacatgagtaacaaaggcatggagcatctgtacagcatgaagtgcaagaacgtggtgccc 2070
          ||   ||||||||| |||||||| ||||||  || |||||||||||| || ||||| ||
Qy   1475 CATGCGAGTAACAAGGGCATGGAACATCTGCTCAACATGAAGTGCAAAAATGTGGTCCCA 1534

Db   2071 ctctatgacctgctgctggagatgctggacgcccac 2106
           | |||||||||||||||||||||||| | ||||||
Qy   1535 GTGTATGACCTGCTGCTGGAGATGCTGAATGCCCAC 1570


RESULT    11
LOCUS       HSERR           6450 bp    RNA             PRI       12-SEP-1993
DEFINITION  Human mRNA for oestrogen receptor.
ACCESSION   X03635 M11457
NID         g31233
KEYWORDS    estrogen receptor; receptor; steroid hormone receptor.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 6450)
  AUTHORS   Green,S., Walter,P., Kumar,V., Krust,A , Bornert,J.M., Argos,P. and
            Chambon,P.
  TITLE     Human oestrogen receptor cDNA: sequence, expression and homology to
            v-erb-A
  JOURNAL   Nature 320 (6058), 134-139 (1986)
  MEDLINE   86146892
COMMENT     Data kindly reviewed (28-OCT-1986) by P. Chambon.
FEATURES             Location/Qualifiers
     source          1..6450
                     /organism="Homo sapiens"
     promoter        26..32
                     /note="CAAT-box"
     promoter        102..109
                     /note="TATA-box"
     misc_RNA        129
                     /note="cap site"
     misc_feature    247..306
                     /note="pot. ORF"
     CDS             361..2148
                     /codon_start=1
                     /product="oestrogen receptor"
                     /db_xref="PID:g31234"
                     /db_xref="SWISS-PROT:P03372"
                     /translation="MTMTLHTKASGMALLHQIQGNELEPLNRPQLKIPLERPLGEVYL
                     DSSKPAVYNYPEGAAYEFNAAAAANAQVYGQTGLPYGPGSEAAAFGSNGLGGFPPLNS
                     VSPSPLMLLHPPPQLSPFLQPHGQQVPYYLENEPSGYTVREAGPPAFYRPNSDNRRQG
                     GRERLASTNDKGSMAMESAKETRYCAVCNDYASGYHYGVWSCEGCKAFFKRSIQGHND
                     YMCPATNQCTIDKNRRKSCQACRLRKCYEVGMMKGGIRKDRRGGRMLKHKRQRDDGEG
                     RGEVGSAGDMRAANLWPSPLMIKRSKKNSLALSLTADQMVSALLDAEPPILYSEYDPT
                     RPFSEASMMGLLTNLADRELVHMINWAKRVPGFVDLTLHDQVHLLECAWLEILMIGLV
                     WRSMEHPVKLLFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEFVCLKS
                     IILLNSGVYTFLSSTLKSLEEKDHIHRVLDKITDTLIHLMAKAGLTLQQQHQRLAQLL
                     LILSHIRHMSNKGMEHLYSMKCKNVVPLYDLLLEMLDAHRLHAPTSRGGASVEETDQS
                     HLATAGSTSSHSLQKYYITGEAEGFPATV"
     misc_feature    442..456
                     /note="pot. nuclear transfer signal"
     misc_feature    1057..1068
                     /note="pot. cAMP-dependent kinase phosphorylation site"
     misc_feature    1264..1275
                     /note="pot. cAMP-dependent kinase phosphorylation site"
     misc_feature    4550..4555
                     /note="pot. alternate polyA signal"
     misc_feature    5625..5630
                     /note="pot. alternate polyA signal"
     polyA_signal    6435..6440
     polyA_site      6450
BASE COUNT     1642 a   1496 c   1534 g   1778 t
ORIGIN

  Query Match             10.4%;  Score 197;  DB 74;  Length 6450;
  Best Local Similarity   66.7%;  Pred. No. 3.65e-171;
  Matches  424;  Conservative    0;  Mismatches  209;  Indels    3;  Gaps    2;

Db    1366 cccttcagtgaagcttcgatgatgggcttactgaccaacctggcagacagggagctggtt 1425
            ||||||  || || || ||||||    |  ||||||||   |||| |||| |||| ||||
Qy     938 CCCTTCACCGAGGCCTCCATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTA 997

Db    1426 cacatgatcaactgggcgaagagggtgccaggctttgtggatttgaccctccatgatcag 1485
           ||||||||| |||||| |||| | | || |||||||||||    | | |||     || ||
Qy     998 CACATGATCAGCTGGGCCAAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAA 1057

Db    1486 gtccaccttctagaatgtgcctggctagagatcctgatgattggtctcgtctggcgctcc 1545
           || |  ||  | ||  |     ||| | ||| |  | ||||| || ||   | ||||||||
Qy    1058 GTGCGGCTCTTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCA 1117

Db    1546 atggagcacccagtgaagctactgtttgctcctaacttgctcttggacaggaaccaggga 1605
           || || ||||| |  ||||| | ||||||||   |   |  |  |||||||| |   ||||
Qy    1118 ATTGACCAGCCCGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGG 1177

Db    1606 aaatgtgtagagggcatggtggagatcttcgacatgctgctggctacatcatctcggttc 1665
           ||||| ||||| || ||  |||| ||||| ||||||||  ||||| ||   | ||  ||||
Qy    1178 AAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTT 1237

Db    1666 cgcatgatgaatctgcagggagaggagtttgtgtgcctcaaatctattatttgcttaat 1725
           |     | | || || ||    | || | | | ||  ||||  | || ||  |||| |||
Qy    1238 CGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAAT 1297

Db    1726 tctggagtgtacacatttctgtccagcaccctgaagtctctggaagagaaggaccatatc 1785
           ||  |  |||||| | | |    |||| | | ||     |  ||  |||   |  |
Qy    1298 TCCAGTATGTACCCTCTGGTCACAGCGACCCAGGATGCTGAC-AGCAGCCGGAAGCTGGC 1355

Db    1786 gaccgagtcctggacaagatcacagacactttgatccacctaatggccaaggcaggcctc 1845
             |  ||| ||    | || ||  ||||| |      |||| ||||||    ||| |
Qy    1357 TCACTTG--CTGAACGCCGTGACCGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATC 1414

Db    1846 acctgcagcagcagcaccggcggctggcccagctgctcctcatcctgtcccacatcagg 1905
            ||   ||||||||    |  ||| |||||  | |||||  |  |||| |||||| |||||
Qy    1415 TCCTCCCAGCAGCAATCCATGCGCCTGGCTAACCTCCTGATGCTCCTGTCCCACGTCAGG 1474

Db    1906 cacatgagtaacaaaggcatggagcatctgtacagcatgaagtgcaagaacgtggtgccc 1965
            ||   ||||||||| |||||||| ||||||  || |||||||||||| || ||||| ||
Qy    1475 CATGCGAGTAACAAGGGCATGGAACATCTGCTCAACATGAAGTGCAAAAATGTGCTCCCA 1534

Db    1966 ctctatgacctgctgctggagatgctggacgcccac 2001
            | |||||||||||||||||||||||||| | ||||||
Qy    1535 GTCTATGACTTGCTGCTGGAGATGCTCAATGCCCAC 1570


RESULT   12
LOCUS       ZFIER        2792 bp    mRNA             VRT       23-JUL-1996
DEFINITION  Poephila guttata estrogen receptor mRNA, complete cds.
ACCESSION   L79911
NID         g1449145
KEYWORDS    estrogen receptor.
SOURCE      Poephila guttata female adult hypothalamus-preoptic area cDNA to
            mRNA.
  ORGANISM  Poephila guttata
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Archosauria; Aves; Neognathae; Passeriformes;
            Estrildidae; Estrildinae; Poephila.
REFERENCE   1  (bases 1 to 2792)
  AUTHORS   Jacobs,E.C., Arnold,A.P. and Campagnoni,A.T.
  TITLE     Zebra finch estrogen receptor cDNA, cloning and mRNA expression
  JOURNAL   J. Steroid Biochem. Mol. Biol. 59(2) (1996) In press
FEATURES             Location/Qualifiers
     source          1..2792
                     /organism="Poephila guttata"
                     /dev_stage="adult"
                     /sex="female"
                     /tissue_type="hypothalamus-preoptic area"
     5'UTR           <1..132
     exon            <1..560
                     /note="Exon boundaries based on alignment to chicken ER"
                     /number=1
     misc_feature    127..237
                     /note="domain A"
     CDS             133..1896
                     /codon_start=1
                     /product="estrogen receptor"
                     /db_xref="PID:g1449146"
                     /translation="MTIHTKTSGVTLLHQIQGTELETLSRPQLKIPLERSLSDMYVET
                     NKTGVFNYPEGATYDFGTTAPVYSSTTLSYAPTSESFGSSSLAGFHSLNSVPPSPVVF
                     LQTAPHWSFFIHHHSQQVPYYLENDQGSFGMREAAPPAFYRPNSDNRRHSIRERMSSA
                     NEKGSLSMESTKETRYCAVCNDYASGYHYGVWSCEGCKAFFKRSIQGHNDYMCPATNQ
                     CTIDKNRRKSCQACRLRKCYEVGMMKGGIRKDRRGGRVMKQKRQRHEQDSRNGEASST
                     ELRAPTLWASPLVVKHNKKNSPALSLTAEQMVSALLEAEPPLVYSEYDPNRPFNEASM
                     MTLLTNLADRELVHMINWAKRVPGFVDLILHDQVHLLECAWLEILMIGLVWRSMEHPG
                     KLLFAPNLLLDRNQGKCVEGMVEIFDMLLATAARFRMMNLQGEEFVCLKSIILLNSGV
                     YTFLSSTLKSLEEKDYIHRVLDKIIDTLIHLMAKSGLSLQQQHRRLAQLLLILSHIRH
                     MSNKGMEHLYNMKCKNVVPLYDLLLEMLDAHRLHAPAARSAAPMEEENRSQLTTASAS
                     SHSLQSFYINSKEEENMQNTL"
     misc_feature    238..645
                     /note="domain B"
     exon            561..751
                     /number=2
     misc_feature    646..894
                     /note="domain C; DNA binding domain"
     exon            752..868
                     /number=3
     exon            869..1204
                     /number=4
     misc_feature    895..1011
                     /note="domain D"
     misc_feature    1012..1764
                     /note="domain E; ligand binding domain"
     exon            1205..1343
```

```
                     /number=5
     exon            1344..1477
                     /number=6
     exon            1478..1661
                     /number=7
     exon            1662..>2792
                     /number=8
     misc_feature    1765..1893
                     /note="domain F"
     3'UTR           1897..>2792
     polyA_signal    2737..2743
BASE COUNT      794 a    669 c    615 g    714 t
ORIGIN

  Query Match             10.0%;  Score 189;  DB 52;  Length 2792;
  Best Local Similarity   66.0%;  Pred. No. 1.48e-162;
  Matches  420;  Conservative    0;  Mismatches  213;  Indels    3;  Gaps    3;

Db   1114 ccattcaatgaagcttctatgatgaccctgttgaccaaccttgcagacagagaattggtg 1173
          || ||||  || || || |||||||      |||||||  | || ||||  || |||||
Qy    938 CCCTTCACCGAGGCCTCCATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTA 997

Db   1174 cacatgatcaactgggcaaagagagttccaggatttgtggatttaacactccatgatcag 1233
          |||||||||| |||||| || |   |||| || ||||||||  | |  ||    || ||
Qy    998 CACATGATCAGCTGGGCCAACAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAA 1057

Db   1234 gtccatctactggaatgtgcatggttagagatactgatgatcggcttagtgtggcgctcc 1293
          || |  || |||||  |    ||| | ||| |  | ||||| ||  |  |||||||||
Qy   1058 GTGCGGCTCCTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCA 1117

Db   1294 atggaacacccaggaaagcttttgtttgcacctaacctgttactggacaggaatcaaggg 1353
          || || ||||| || ||||||| | ||||||||  |  |  |  |||||||| || | |||
Qy   1118 ATTGACCACCCCGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGG 1177

Db   1354 aaatgtgtagagggcatggtggaaatctttgacatgctgctggctactgctgctcggttt 1413
          ||||| ||||| || ||  ||||||||||||||||||| ||||| ||| || |  |||||
Qy   1178 AAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTT 1237

Db   1414 cgcatgatgaaccttcaaggggaggaatttgtgtgccttaagtccatcatcctgctcaat 1473
          ||   | | || || |||    | |||| | | ||  | ||| |||| ||||||||||||
Qy   1238 CGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAAT 1297

Db   1474 tctggtgtgtacacttttctttccagcaccttgaaatctctggaagagaaagactatatt 1533
          ||  || ||||| || |  |  |    |||  | |  ||    |  ||   ||   |
Qy   1298 TCCAGTATGTACCCTCTGGTCACAGCGACCCAGGATGCTGAC-AGCAGCCGGAAGCTGG- 1355

Db   1534 caccgtgtcctggataaaatcacagatactctgatacacttaatggctaagtcaggtctt 1593
          | |  | | || |     | || ||| || || |     | || || |||   ||  |
Qy   1356 CTCACT-TGCTGAACGCCGTGACCGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATC 1414

Db   1594 tctctgcagcagcaacacagacgactggctcagctcctcctcatcctctctcatatcaga 1653
          ||    |||||||||  ||  || |||||| | |||||  |  |||| || ||  ||||
Qy   1415 TCCTCCCAGCAGCAATCCATGCGCCTGGCTAACCTCCTGATGCTCCTGTCCCACGTCAGG 1474

Db   1654 cacatgagcaacaaaggaatggagcacctgtacaatatgaagtgtaagaatgtagttccg 1713
          ||   ||| ||||| || ||||| || |||  ||| ||||||||| || ||||| || ||
Qy   1475 CATGCGAGTAACAAGGGCATGGAACATCTGCTCAACATGAAGTGCAAAAATGTGGTCCCA 1534

Db   1714 ctctacgatctcttgctggagatgctggatgctcac 1749
           | || || ||  ||||||||||||||| |||| |||
Qy   1535 GTGTATGACCTGCTGCTGGAGATGCTGAATGCCCAC 1570


RESULT      13
LOCUS       RNESTROR     1880 bp    RNA             ROD       29-OCT-1996
DEFINITION  Rat mRNA for estrogen receptor.
ACCESSION   X61098
NID         g56120
KEYWORDS    estrogen receptor; steroid hormone receptor.
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 1880)
  AUTHORS   Maggi,A.M.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-JUN-1991) A M.A. Maggi, Milano Mol Pharm Lab, Inst
            Pharmacological Sciences, Piazza Durante 11, 20131 Milano, ITALY
REFERENCE   2  (bases 1 to 1880)
  AUTHORS   Spreafico,E., Bettini,E., Pollio,G. and Maggi,A.
  TITLE     Nucleotide sequence of estrogen receptor cDNA from Sprague-Dawley
            rat
  JOURNAL   Eur. J. Pharmacol. 227 (3), 353-356 (1992)
  MEDLINE   93114361
FEATURES             Location/Qualifiers
     source          1..1880
                     /organism="Rattus norvegicus"
                     /strain="Sprague-Dawley"
                     /tissue_type="uterus"
     mRNA            1..1880
                     /evidence=experimental
                     /note="estrogen receptor"
     CDS             74..1876
                     /codon_start=1
                     /product="estrogen receeptor"
                     /db_xref="PID:g56121"
                     /db_xref="SWISS-PROT:P06211"
                     /translation="MTMTLHTKASGMALLHQIQGNELEPLNRPQLKMPMERALGEVYV
                     DNSKPAVFNYPEGAAYEFNAAAAAAAAGASAPVYGQSSITYGPGSEAAAFGANSLGAF
                     PQLNSVSPSPLMLLHPPPHVSPFLHPHGHQVPYYLENEPSAYAVRDIGPPAFYRSNSD
                     NRRQNGRERLSSSSEKGNMIMESAKETRYCAVCNDYASGYHYGVWSCEGCKAFFKRSI
                     QGHNDYMCPATNQCTIDKNRRKSCQACRLRKCYEVGMMKGGIRKDRRGGRMLKHKRQR
                     DDLEGRNEMGTSGDMRAANLWPSPLVIKHTKKNSPALSLTADQMVSALLDAEPPLIYS
                     EYDPSRPFSEASMMGLLTNLADRELVHMINWAKRVPGFGDLNLHDQVHLLECAWLEIL
                     MIGLVWRSMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEF
                     VCLKSIILLNSGVYTFLSSTLKSLEEKDHIHRVLDKITDTLIHLMAKAGLTLQQQHRR
                     LAQLLLILSHIRHMSNKGMEHLYNMKCKNVVPLYDLLLEMLDAHRLHAPASRMGVPPE
                     EPSQSQLTTTSSTSAHSLQTYYIPPEAEGFPNTI"
BASE COUNT      438 a    548 c    505 g    389 t
ORIGIN

  Query Match              9.7%;  Score 184;  DB 91;  Length 1880;
  Best Local Similarity   65.7%;  Pred. No. 3.46e-157;
  Matches  417;  Conservative    0;  Mismatches  215;  Indels    3;  Gaps    3;

Db   1094 cccttcagtgaagcctcaatgatgggcttattgaccaacctggcagacagggagctggtt 1153
          |||||||  || ||||| ||||||   |   |||||||  |||| |||| |||| ||||
Qy    938 CCCTTCACCGAGGCCTCCATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTA 997

Db   1154 catatgatcaactgggcaaagagagtgccaggctttggggacttgaatctccacgatcaa 1213
          || ||||||| |||||| || |   | || ||||||| |||  | |  ||   ||| |||
Qy    998 CACATGATCAGCTGGGCCAACAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAA 1057

Db   1214 gttcaccttctggagtgtgcctggttggagatcctgatgattggtctggtctggcgctcc 1273
          || |  || |||||| |    ||| ||||| |  | ||||| || ||| | ||||||||
Qy   1058 GTGCGGCTCCTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCA 1117

Db   1274 atggaacacccggggaagctcctgtttgctcctaacttgctcttggacaggaatcaaggt 1333
          || || ||||| || ||||||| | ||||||||  |  |  |  |||||||| || | ||
Qy   1118 ATTGACCACCCCGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGG 1177

Db   1334 aaatgtgtagaaggcatggtggagatctttgacatgttgctggctacgtcaagtcgattc 1393
          ||||| |||||||| ||  |||| ||||||||||||| | ||||| ||  |     | ||
Qy   1178 AAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTT 1237

Db   1394 cgcatgatgaacctgcagggagaagagtttgtgtgcctcaaatcaatcattttgcttaat 1453
          ||   | | || || ||     |||| | | | ||  ||||  | || ||  |||| |||
Qy   1238 CGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAAT 1297

Db   1454 tctggagtgtacacatttctatccagcaccttgaagtctctggaagagaaggaccacatc 1513
          ||  |  ||||| |  |  |  |    |||  | |  ||    |  ||  |||   |
Qy   1298 TCCAGTATGTACCCTCTGGTCACAGCGACCCAGGATGCTGAC-AGCAGCCGGAAG-CTGG 1355

Db   1514 caccgagtcctggacaagatcacagacactttgatccacttgatggccaaagctggcctg 1573
          | |    | ||| ||    | || ||  ||||| |     |||| |||||    |||  |
Qy   1356 CTCACT-TGCTGAACGCCGTGACCGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATC 1414

Db   1574 actctgcagcaacagcatcgacgtctggcccagctcctcctcatcctttcccatatccgg 1633
           |    ||||| ||       || |||||  | |||||  |  |||| |||||  || ||
Qy   1415 TCCTCCCAGCAGCAATCCATGCGCCTGGCTAACCTCCTGATGCTCCTGTCCCACGTCAGG 1474

Db   1634 cacatgagtaacaaaggcatggagcatctctacaacatgaaatgcaagaatgtcgtgcct 1693
          ||   ||||||||| |||||||| |||||   ||||||||| ||||| ||||| || ||
Qy   1475 CATGCGAGTAACAAGGGCATGGAACATCTGCTCAACATGAAGTGCAAAAATGTGGTCCCA 1534

Db   1694 ctctatgacctgctgctggagatgctggatgctca 1728
           | |||||||||||||||||||||||| |||| ||
Qy   1535 GTGTATGACCTGCTGCTGGAGATGCTGAATGCCCA 1569


RESULT    14
LOCUS       RNERR        2090 bp    RNA             ROD       12-SEP-1993
DEFINITION  Rat mRNA for estrogen receptor.
ACCESSION   Y00102
NID         g56110
KEYWORDS    estrogen receptor
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 2090)
  AUTHORS   Muramatsu,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-MAR-1987) Masami Muramatsu,Dept of Biochemistry,
            Faculty of Medicine The University of Tokyo, Hongo 7-3-1,
            Bunkyo-Ku, Tokyo 113 Japan
REFERENCE   2  (bases 1 to 2090)
  AUTHORS   Koike,S., Sakai,M. and Muramatsu,M.
  TITLE     Molecular cloning and characterization of rat estrogen receptor
            cDNA
  JOURNAL   Nucleic Acids Res. 15 (6), 2499-2513 (1987)
  MEDLINE   87174780
FEATURES             Location/Qualifiers
     source          1..2090
                     /organism="Rattus norvegicus"
                     /strain="Wistar"
                     /tissue_type="utems"
                     /clone_lib="rat uterus cDNA 1."
     CDS             210..2012
                     /note="estrogen receptor protein (AA 1-600)"
                     /codon_start=1
                     /db_xref="PID:g56111"
                     /db_xref="SWISS-PROT:P06211"
                     /translation="MTMTLHTKASGMALLHQIQGNELEPLNRPQLKMPMERALGEVYV
                     DNSKPAVFNYPEGAAYEFNAAAAAAAAGASAPVYGQSSITYGPGSEAAAFGANSLGAF
                     PQLNSVSPSPLMLLHPPPHVSPFLHPHGHQVPYYLENEPSAYAVRDTGPPAFYRSNSD
                     NRRQNGRERLSSSSEKGNMIMESAKETRYCAVCNDYASGYHYGVWSCEGCKAFFKRSI
                     QGHNDYMCPATNQCTIDKNRRKSCQACRLRKCYEVGMMKGGIRKDRRGGRMLKHKRQR
                     DDLEGRNEMGTSGDMRAANLWPSPLVIKHTKKNSPALSLTADQMVSALLDAEPPLIYS
                     EYDPSRPFSEASMMGLLTNLADRELVHMINWAKRVPGFGDLNLHDQVHLLECAWLEIL
                     MIGLVWRSMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEF
                     VCLKSIILLNSGVYTFLSSTLKSLEEKDHIHRVLDKINDTLIHLMAKAGLTLQQQHRR
                     LAQLLLILSHIRHMSNKGMEHLYNMKCKNVVPLYDLLLEMLDAHRLHAPASRMGVPPE
                     EPSQSQLTTTSSTSAHSLQTYYIPPEAEGFPNTI"
BASE COUNT      488 a    619 c    537 g    446 t
ORIGIN

  Query Match             9.7%;  Score 184;  DB 91;  Length 2090;
  Best Local Similarity  65.7%;  Pred. No. 3.46e-157;
  Matches  417;  Conservative    0;  Mismatches 215;  Indels    3;  Gaps    3;

Db   1230 cccttcagtgaagcctcaatgatgggcttattgaccaacctggcagacagggagctggtt 1289
          |||||||  || ||||| ||||||   |   |||||||  |||| |||| |||| ||||
Qy    938 CCCTTCACCGAGGCCTCCATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTA 997

Db   1290 catatgatcaactgggcaaagagagtgccaggctttggggacttgaatctccacgatcaa 1349
          || ||||||| |||||| ||||   | || |||||||| |||  | |  ||   ||| |||
Qy    998 CACATGATCAGCTGGGCCAAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAA 1057

Db   1350 gttcaccttctggagtgtgcctggttggagatcctgatgattggtctggtctggcgctcc 1409
          || |  ||  ||||| |    ||| ||||| |  | ||||| || ||| | ||||||||
Qy   1058 GTGCGGCTCTTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCA 1117

Db   1410 atggaacacccggggaagctcctgtttgctcctaacttgctcttggacaggaatcaaggt 1469
          || || ||||| || |||||| | ||||||||  |  |  |  |||||||| || | ||
Qy   1118 ATTGACCACCCCGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGG 1177

Db   1470 aaatgtgtagaaggcatggtggagatctttgacatgttgctggctacgtcaagtcgattc 1529
          ||||| |||||||| ||  |||| |||||||||||| | ||||| ||  |     | ||
Qy   1178 AAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTT 1237

Db   1530 cgcatgatgaacctgcagggagaagagtttgtgtgcctcaaatcaatcattttgcttaat 1589
          ||   | | || || ||     |||| | | | ||  ||||  | || ||  |||| |||
Qy   1238 CGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAAT 1297

Db   1590 tctggagtgtacacatttctatccagcaccttgaagtctctggaagagaaggaccacatc 1649
          ||  |  ||||| |  |  |  |    |||  | |  ||    |  ||  |||   |
Qy   1298 TCCAGTATGTACCCTCTGGTCACAGCGACCCAGGATGCTGAC-AGCAGCCGGAAG-CTGG 1355

Db   1650 caccgagtcctggacaagatcaatgacactttgatccacttgatggccaaagctggcctg 1709
          | |    | ||| ||    | |  ||  ||||| |     |||| |||||    |
Qy   1356 CTCACT-TGCTGAACGCCGTGACCGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATC 1414

Db   1710 actctgcagcaacagcatcgacgtctggcccagctcctcctcatcctttcccatatccgg 1769
           |    ||||| ||       || |||||  | |||||  |  |||| |||||  || ||
Qy   1415 TCCTCCCAGCAGCAATCCATGCGCCTGGCTAACCTCCTGATGCTCCTGTCCCACGTCAGG 1474

Db   1770 cacatgagtaacaaaggcatggagcatctctacaacatgaaatgcaagaatgtcgtgcct 1829
          ||   ||||||||| |||||||| |||||   ||||||||| ||||| ||||| || ||
Qy   1475 CATGCGAGTAACAAGGGCATGGAACATCTGCTCAACATGAAGTGCAAAAATGTGGTCCCA 1534

Db   1830 ctctatgacctgctgctggagatgctggatgctca 1864
           | |||||||||||||||||||||||| |||| ||
Qy   1535 GTGTATGACCTGCTGCTGGAGATGCTGAATGCCCA 1569


RESULT    15
LOCUS       S52128       2090 bp    mRNA            ROD       12-MAR-1993
DEFINITION  estrogen receptor [rats, brain, mRNA, 2090 nt].
ACCESSION   S52128
NID         g262708
KEYWORDS    .
SOURCE      Rattus sp. brain.
  ORGANISM  Rattus sp.
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 2090)
  AUTHORS   Bettini,E., Pollio,G., Santagati,S. and Maggi,A.
  TITLE     Estrogen receptor in rat brain: presence in the hippocampal
            formation
  JOURNAL   Neuroendocrinology 56 (4), 502-508 (1992)
  MEDLINE   93116840
  REMARK    GenBank staff at the National Library of Medicine created this
            entry [NCBI gibbsq 121564] from the original journal article.
            This sequence comes from Fig. 1.
FEATURES             Location/Qualifiers
     source          1..2090
                     /organism="Rattus sp."
BASE COUNT      488 a    619 c    537 g    446 t
ORIGIN

  Query Match             9.7%;  Score 184;  DB 92;  Length 2090;
  Best Local Similarity  65.7%;  Pred. No. 3.46e-157;
  Matches  417;  Conservative    0;  Mismatches 215;  Indels    3;  Gaps    3;

Db   1230 cccttcagtgaagcctcaatgatgggcttattgaccaacctggcagacagggagctggtt 1289
          |||||||  || ||||| ||||||   |   |||||||  |||| |||| |||| ||||
Qy    938 CCCTTCACCGAGGCCTCCATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTA 997
```

```
Db   1290 catatgatcaactgggcaaagagagtgccaggctttggggacttgaatctccacgatcaa 1349
          || ||||||| |||||| ||||   | || ||||||| |||  | |  ||   ||| |||
Qy    998 CACATGATCAGCTGGGCCAAGAACATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAA 1057

Db   1350 gttcaccttctggagtgtgcctggttggagatcctgatgattggtctggtctggcgctcc 1409
          || |  ||  ||||| |    ||| ||||| |  | ||||| || ||| | ||||||||
Qy   1058 GTGCGGCTCTTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCA 1117

Db   1410 atggaacacccggggaagctcctctttactcctaacttgctcttggacaggaatcaaggt 1469
          || || ||| | || |||||| | ||||||||  |  |  |  |||||||| || | ||
Qy   1118 ATTGACCACCCCGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGG 1177

Db   1470 aaatgtgtagaaggcatggtggagatctttgacatgttgctggctacgtcaagtcgattc 1529
          ||||| ||| |||| ||  |||| |||||||||||| | ||||| ||  |     | ||
Qy   1178 AAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTT 1237

Db   1530 cgcatgatgaacctgcagggagaagagtttgtgtgcctcaaatcaatcattttgcttaat 1589
          ||   | | || || ||     |||| | | | ||  ||||  | || ||  |||| |||
Qy   1238 CGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAAT 1297

Db   1590 tctggagtgtacacatttctatccagcaccttgaagtctctggaagagaaggaccacatc 1649
          ||  |  ||||| |  |  |  |    |||  | |  ||    |  ||  |||   |
Qy   1298 TCCAGTATGTACCCTCTGGTCACAGCGACCCAGGATGCTGAC-AGCAGCGGAAG-CTGG 1355

Db   1650 caccgagtcctggacaagatcaacgacactttgatccacttgatggccaaagctggcctg 1709
          | |    | ||| ||    | | |||  ||||| |     |||| |||||    ||| |
Qy   1356 CTCACT-TGCTGAACGCCGTGACCGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATC 1414

Db   1710 actctgcagcaacagcatcggcgtctggccagctcctcctcatcctttcccatatccgg 1769
           |    ||||| ||       || |||||  | |||||  |  |||| |||||  || ||
Qy   1415 TCCTCCCAGCAGCAATCCATGGGCCTGGCTAACCTCCTGATGCTCCTGTCCCACGTCAGG 1474

Db   1770 caatgagtaacaaggcatggaacatctctacaacatgaaatgcaagaatgtcgtgcct 1829
          ||   ||||||||| |||||||| |||||   ||||||||| ||||| ||||| || ||
Qy   1475 CATGCGAGTAACAAGGGCATGGAACATCTGCTCAACATGAAGTGCAAAAATGTGGTCCCA 1534

Db   1830 ctctatgacctgctgctggagatgctggatgctca 1864
           | |||||||||||||||||||||||| |||| ||
Qy   1535 GTGTATGACCTGCTGCTGGAGATGCTGAATGCCCA 1569

Search completed: Thu Dec 18 20:47:03 1997
Job time : 1604 secs.
```

```
***************************************************************************
  ___   ___    ______     ______     ______     ______     _    _
 |   \ /   |  |  __  |   |  ____|   |  __  |   |  ____|   | |  | |
 | |\ V /| |  | |__| |   | |____    | |__| |   | |        | |__| |
 | | \_/ | |  |  ____|   |____  |   |  _  /    | |        |  __  |
 | |     | |  | |         ____| |   | | \ \    | |____    | |  | |
 |_|     |_|  |_|        |______|   |_|  \_\   |______|   |_|  |_|  (TM)

***************************************************************************

          Release 2.1D John F. Collins, Biocomputing Research Unit
          Copyright (c) 1993, 1994, 1995  University of Edinburgh, U.K.
                  Distribution rights by IntelliGenetics, Inc

MPsrch_nn   n.a  -  n.a  database search, using Smith-Waterman algorithm

Run on:           Thu Dec 18 20:47:23 1997;  MasPar time 207.93 Seconds
                                             952.646 Million cell updates/sec
Tabular output not generated.

Title:            >US-08-826-361A-24
Description       (1-1898) from US08826361A.seq
Perfect Score     1898
N.A. Sequence        1 CACGAATCTTTGAGAACATT        ...TGAATGGAGGGTTTTTACAC 1898
     Comp:               GTGCTTAGAAACTCTTGTAA          ACTTACCTCCCAAAGGTGTG

Scoring table:    TABLE default
                  Gap 6

Nmatch   STD :    Dbase 0;  Query 0

Searched:         142090 seqs, 52183452 bases x 2

Post-processing: Minimum Match 0%
                  Listing first 45 summaries

Database:         n-geneseq28
                    1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
                    8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
                    14:part14 15:part15 16:part16 17:part17 18:part18
                    19:part19 20:part20 21:part21 22:part22 23:part23
                    24:part24 25:part25 26:part26 27:part27 28:part28
                    29:part29

Statistics:       Mean 9.610;  Variance 5.192;  scale 1.851

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution

                                 SUMMARIES
                %
Result          Query
   No   Score   Match Length DB  ID               Description             Pred. No.
-----------------------------------------------------------------------------------
     1   1452    76.5   1460 29  T62843           Human oestrogen recep  0.00e+00
     2   1139    60.0   2568 29  T62842           Rat oestrogen recepto  0.00e+00
     3   1010    53.2   1459 29  T62944           Mouse oestrogen recep  0.00e+00
     4    197    10.4   4963 17  T05873           RhCMV*-1 promoter and  6.70e-125
     5    197    10.4   4963 13  Q76272           RhCMV*-1-controlled p  6.70e-125
     6    195    10.3   2092  2  N70980           cDNA encoding human t  2.29e-123
     7    127     6.7    204 14  Q85976           Human estrogen recept  8.12e-72
     8     91     4.8   1047  2  Q10572           Human Natriuretic Pep  2.71e-45
     9     75     4.0   1047  2  Q10572           Human Natriuretic Pep  7.88e-34
    10     71     3.7   3460 24  T42193           Fragment of pig oestr  5.20e-31
    11     65     3.4   2157  2  N80021           Sequence encoding hum  5.54e-29
    12     63     3.3   6244 13  Q76269           RhCMV*-1-controlled p  1.92e-25
    13     63     3.3   6244 17  T05870           RhCMV*-1 promoter and  1.92e-25
    14     58     3.1   1912  2  Q10388           Encodes skin-specific  5.08e-22
    15     58     3.1   2740  3  Q20542           Mouse retinoic acid r  5.08e-22
    16     58     3.1   2740  3  Q20541           Mouse retinoic acid r  5.08e-22
    17     59     3.1   2992 13  Q79934           Human liver hap cDNA.  1.06e-22
    18     59     3.1   2992 11  Q65572           Human liver hap cDNA   1.06e-22
    19     59     3.1   2992 20  T05491           Human hepatoma retino  1.06e-22
    20     58     3.1   5040  2  N80916           Sequence encoding the  5.08e-22
    21     56     3.0   1913 18  T05202           Murine retinoic acid   1.15e-20
    22     56     3.0   6176 27  T49220           Plasmid pGR0403R for   1.15e-20
    23     55     2.9   2999  1  N90093           Human Hap (hepatoma)   5.43e-20
    24     53     2.8   1231  6  Q37761           Sequence of a 1.23 kb  1.19e-18
    25     53     2.8   1785  2  Q12005           Human TR2-9 DNA bindi  1.19e-18
    26     53     2.8   2029  2  Q12004           Human TR2-5 androgen   1.19e-18
    27     53     2.8   2221  2  Q12006           Human TR2-11 DNA bind  1.19e-18
    28     53     2.8   2458  2  Q12003           Human TR2-7 DNA bindi  1.19e-18
    29     51     2.7   1576  2  Q10226           Clone pGEM-hRARgamma   2.55e-17
    30     51     2.7   1903  2  Q10389           Encodes human Retinoi  2.55e-17
    31     51     2.7   1903 18  T05196           Human retinoic acid r  2.55e-17
    32     51     2.7   3569  1  N91772           Human androgen recept  2.55e-17
    33     51     2.7   3569 29  T63407           Androgen receptor cDN  2.55e-17
    34     51     2.7   3715  2  Q12001           Full-length human and  2.55e-17
    35     49     2.6    204  1  N81164           Base substituted F co  5.33e-16
    36     49     2.6    204  1  N81164           Base substituted F co  5.33e-16
    37     50     2.6   1351 13  Q79935           Human liver RAR-beta   1.17e-16
    38     50     2.6   2249  2  N80021           Sequence encoding hum  1.17e-16
    39     47     2.5    387 24  T45454           Modified human glucoc  1.08e-14
    40     47     2.5    387 19  T09863           Modified human glucoc  1.08e-14
    41     47     2.5   3217  1  N91578           Rat androgen receptor  1.08e-14
    42     45     2.4   1893  2  Q12008           TrpE/androgen recepto  2.14e-13
    43     45     2.4   3217  2  Q12002           Full-length rat andro  2.14e-13
    44     45     2.4   3427 10  Q56733           cDNA sequence encodin  2.14e-13
    45     45     2.4   4190  1  N91773           Rat androgen receptor  2.14e-13

                                ALIGNMENTS

RESULT    1
ID   T62843 standard; cDNA; 1460 BP.
AC   T62843;
DT   08-JUN-1997  (first entry)
DE   Human oestrogen receptor beta (ER-beta) cDNA.
KW   Orphan receptor; oestrogen receptor beta; ER-beta;
KW   nuclear receptor; prostate cancer; benign prostatic hyperplasia;
KW   ovary cancer; cardiovascular disease; osteoporosis;
KW   environment; pollutant; ss.
OS   Homo sapiens.
FH   Key                 Location/Qualifiers
FT   CDS                 3..1460
FT   /*tag=  a
PN   WO9709348-A2.
PD   13-MAR-1997.
PF   09-SEP-1996; E03933.
PR   08-SEP-1995; GB-018272.
PR   15-MAR-1996; GB-005550.
PR   11-APR-1996; GB-007532.
PR   08-MAY-1996; GB-009576.
PA   (KARO-) KARO BIO AB.
PI   Enmark E,  Gustafsson J,  Kuiper GG;
DR   WPI; 97-192842/17.
DR   P-PSDB; W14724.
PT   New isolated oestrogen receptor beta - used to develop prods. for
PT   treating e.g. cancers, CNS diseases, osteoporosis or cardiovascular
PT   disease
PS   Claim 6; Fig 13B; 45pp; English.
CC   A cDNA clone (T62843) codes for a novel human oestrogen receptor-
CC   related nuclear receptor (W14724), designated ER-beta.  It was
CC   isolated from a human testis cDNA library using rat ER-beta cDNA
CC   (see also T62842) as probe.  The human ER-beta gene was mapped
CC   to chromosome 14 region 14q22-23.  Rat, human and mouse ER-beta
CC   (W14723-25) can be used to isolate molecules for use in the
CC   treatment of cardiovascular diseases, central nervous system
CC   diseases, osteoporosis, prostate or ovarian cancer or benign
CC   prostatic hyperplasia and to test environmental chemicals for
CC   oestrogenic activity.
SQ   Sequence   1460 BP;    352 A;    379 C;    411 G;    318 T;
```

```
  Query Match             76.5%;  Score 1452;  DB 29;  Length 1460;
  Best Local Similarity   99.8%;  Pred. No. 0.00e+00;
  Matches 1455;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Db          3 atgacattctacagtcctgctgtgatgaattacagcattcccagcaatgtcactaacttg 62
              ||||||||||| || |||||||||||||||||||||||||||||||||||||||||||||
Qy        212 ATGACATTCTATACCCCTGCTGTGATGAATTACAGCATTCCCAGCAATGTCACTAACTTG 271

Db         63 gaaggtgggcctggtcggcagaccacaagcccaaatgtgttgtggccaacacctgggcac 122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        272 GAAGGTGGGCCTGGTCGGCAGACCACAAGCCCAAATGTGTTGTGGCCAACACCTGGGCAC 331

Db        123 ctttctcctttagtggtccatcgccagttatcacatctgtatgcggaacctcaaaagagt 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        332 CTTTCTCCTTTAGTGGTCCATCGCCAGTTATCACATCTGTATGCGGAACCTCAAAAGAGT 391

Db        183 ccctggtgtgaagcaagatcgctagaacacaccttacctgtaaacagagagacactgaaa 242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        392 CCCTGGTGTGAAGCAAGATCGCTAGAACACACCTTACCTGTAAACAGAGAGACACTGAAA 451

Db        243 aggaaggttagtgggaaccgttgcgccagccctgttactggtccaggttcaaagagggat 302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        452 AGGAAGGTTAGTGGGAACCGTTGCGCCAGCCCTGTTACTGGTCCAGGTTCAAAGAGGGAT 511

Db        303 gctcacttctgcgctgtctgcagcgattacgcatcgggatatcactatggagtctggtcg 362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        512 GCTCACTTCTGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCG 571

Db        363 tgtgaaggatgtaaggccttttttaaaagaagcattcaaggacataatgattatatttgt 422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        572 TGTGAAGGATGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGT 631

Db        423 ccagctacaaatcagtgtacaatcgataaaaaccggcgcaagagctgccaggcctgccga 482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        632 CCAGCTACAAATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGA 691

Db        483 cttcggaagtgttacgaagtgggaatggtcaagtgtggctcccggagagagagatgtggg 542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        692 CTTCGGAAGTGTTACGAAGTGGGAATGGTCAAGTGTGGCTCCCGGAGAGAGAGATGTGGG 751

Db        543 taccgccttgtgcggagacagagaagtgccgacgagcagctgcactgtgccggcaaggcc 602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        752 TACCGCCTTGTGCGGAGACAGAGAAGTGCCGACGAGCAGCTGCACTGTGCCGGCAAGGCC 811

Db        603 aagagaagtggcggccacgcgccccgagtgcgggagctgctgctggacgccctgagcccc 662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        812 AAGAGAAGTGGCGGCCACGCGCCCCGAGTGCGGGAGCTGCTGCTGGACGCCCTGAGCCCC 871

Db        663 gagcagctagtgctcaccctcctggaggctgagccgccccatgtgctgatcagccgcccc 722
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        872 GAGCAGCTAGTGCTCACCCTCCTGGAGGCTGAGCCGCCCCATGTGCTGATCAGCCGCCCC 931

Db        723 agtgcgcccttcaccgaggcctccatgatgatgtccctgaccaagttggccgacaaggag 782
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        932 AGTGCGCCCTTCACCGAGGCCTCCATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAG 991

Db        783 ttggtacacatgatcagctgggccaagaagattcccggctttgtggagctcagcctgttc 842
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        992 TTGGTACACATGATCAGCTGGGCCAAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTC 1051

Db        843 gaccaagtgcggctcttggagagctgttggatggaggtgttaatgatggggctgatgtgg 902
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1052 GACCAAGTGCGGCTCTTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGG 1111

Db        903 cgctcaattgaccaccccggcaagctcatctttgctccagatcttgttctggacagggat 962
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1112 CGCTCAATTGACCACCCCGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGAT 1171

Db        963 gaggggaaatgcgtagaaggaattctggaaatctttgacatgctcctggcaactacttca 1022
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1172 GAGGGGAAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTACTTCA 1231

Db       1023 aggtttcgagagttaaaactccaacacaaagaatatctctgtgtcaaggccatgatcctg 1082
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1232 AGGTTTCGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGATCCTG 1291

Db       1083 ctcaattccagtatgtaccctctggtcacagcgacccaggatgctgacagcagccggaag 1142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1292 CTCAATTCCAGTATGTACCCTCTGGTCACAGCGACCCAGGATGCTGACAGCAGCCGGAAG 1351

Db       1143 ctggctcacttgctgaacgccgtgaccgatgctttggtttgggtgattgccaagagcggc 1202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1352 CTGGCTCACTTGCTGAACGCCGTGACCGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGC 1411

Db       1203 atctcctcccagcagcaatccatgcgcctggctaacctcctgatgctcctgtcccacgtc 1262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1412 ATCTCCTCCCAGCAGCAATCCATGCGCCTGGCTAACCTCCTGATGCTCCTGTCCCACGTC 1471

Db       1263 aggcatgcgagtaacaagggcatggaacatctgctcaacatgaagtgcaaaaatgtggtc 1322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1472 AGGCATGCGAGTAACAAGGGCATGGAACATCTGCTCAACATGAAGTGCAAAAATGTGGTC 1531

Db       1323 ccagtgtatgacctgctgctggagatgctgaatgcccacgtgcttcgcgggtgcaagtcc 1382
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1532 CCAGTGTATGACCTGCTGCTGGAGATGCTGAATGCCCACGTGCTTCGCGGGTGCAAGTCC 1591

Db       1383 tccatcacggggtccgagtgcagcccggcagaggacagtaaaagcaaagagggctcccag 1442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1592 TCCATCACGGGGTCCGAGTGCAGCCCGGCAGAGGACAGTAAAAGCAAAGAGGGCTCCCAG 1651

Db       1443 aacctacagtctcagtga 1460
              |||| |||||||||||||
Qy       1652 AACCCACAGTCTCAGTGA 1669


RESULT    2
ID   T62842 standard; cDNA; 2568 BP.
AC   T62842;
DT   08-JUN-1997  (first entry)
DE   Rat oestrogen receptor beta (ER-beta) cDNA.
KW   Orphan receptor; oestrogen receptor beta; ER-beta;
KW   nuclear receptor; prostate cancer; benign prostatic hyperplasia;
KW   ovary cancer; cardiovascular disease; osteoporosis;
KW   environment; pollutant; ss.
OS   Rattus sp.
FH   Key             Location/Qualifiers
FT   CDS             424..1881
FT   /*tag=  a
FT   /note= "alternative start codon at 448..450"
PN   WO9709348-A2.
PD   13-MAR-1997.
PF   09-SEP-1996; E03933.
PR   08-SEP-1995; GB-018272.
PR   15-MAR-1996; GB-005559.
PR   11-APR-1996; GB-007532.
PR   08-MAY-1996; GB-009576.
PA   (KARO-) KARO BIO AB.
PI   Enmark E,  Gustafsson J,  Kuiper GG;
DR   WPI; 97-192842/17.
DR   P-PSDB; W14723.
PT   New isolated oestrogen receptor beta - used to develop prods. for
PT   treating e.g. cancers, CNS diseases, osteoporosis or cardiovascular
PT   disease
PS   Claim 6; Fig 1; 45pp; English.
CC   A cDNA clone (T62842) codes for a novel rat oestrogen receptor-
CC   related nuclear receptor (W14723), designated ER-beta.  A set of
CC   degenerate primers based on conserved sequences of the DNA-
CC   binding domain and ligand binding domain of nuclear receptors were
CC   used to amplify rat prostate cDNA.  The PCR prod. was used to
CC   screen a rat prostate cDNA library to obtain the ER-beta cDNA.
CC   This was then used to obtain human ER-beta cDNA (see also T62843)
CC   from a human testis cDNA library.  Rat, human and mouse ER-beta
CC   (W14723-25) can be used to isolate molecules for use in the
CC   treatment of cardiovascular diseases, central nervous system
CC   diseases, osteoporosis, prostate or ovarian cancer or benign
CC   prostatic hyperplasia and to test environmental chemicals for
CC   oestrogenic activity.
SQ   Sequence   2568 BP;    634 A;    654 C;    663 G;    617 T;

  Query Match             60.0%;  Score 1138;  DB 29;  Length 2568;
  Best Local Similarity   84.8%;  Pred. No. 0.00e+00;
  Matches  1418;  Conservative    0;  Mismatches  250;  Indels    5;  Gaps    4;

Db    230 ataatgtccatctgtacctcttctccaaggaggttttcttcagctgcgacctctgaagac 289
          |||||| || | ||| ||||||||   | |  |||||||          | ||  |||||
Qy     21 ATAATGACCTT-TGTGCCTCTTCTTGCAAGGTGTTTTCTCAGCTGTTATC-TC--AAGAC 76

Db    290 atggagatcaaaaactcaccgtcgagcctcagttc-cctgcttcctataactgtagccag 348
          ||||| || ||||||||||| || ||||| | ||| ||| | ||||| ||||| || || 
Qy     77 ATGGATATAAAAAACTCACCATCTAGCCTTAATTCTCCTTCCTCCTACAACTGCAGTCAA 136

Db    349 tccatcctacccctggagcacggctgcatctacatcccttcctcctacgtagacaaccgc 408
          |||||| |||||||||||||||||||||| ||||| |||||||||| ||||||| || |
Qy    137 TCCATCTTACCCCTGGAGCACGGCTGCATATACATACCTTCCTCCTATGTAGACAGCCAC 196

Db    409 catgagtattcagctatgacattctacagtcctgctgtgatgaactacagtgttcccggc 468
          ||||| ||| |||| |||||||||||| || |||||||||||||| |||||  ||||| ||
Qy    197 CATGAATATCCAGCCATGACATTCTATAGCCCTGCTGTGATGAATTACAGCATTCCCAGC 256

Db    469 agcaccagtaacctggacggtgggcctgtccgactgagcacaagcccaaatgtgctatgg 528
          |    || |||| |||| ||||||||||  || | || ||||||||||||||| | |||
Qy    257 AATGTCACTAACTTGGAAGGTGGGCCTGGTCGGCAGACCACAAGCCCAAATGTGTTGTGG 316

Db    529 ccaacttctgggcacctgtctcctttagcgacccattgccaatcatcgctcctctatgca 588
          |||||  |||||||||| |||||||||| |  |||| |||| | ||| |  || ||||| 
Qy    317 CCAACACCTGGGCACCTTTCTCCTTTAGTGGTCCATCGCCAGTTATCACATCTGTATGCG 376

Db    589 gaacctcaaaagagtccctggtgtgaagcaagatcactagagcacaccttacctgtaaac 648
          ||||||||||||||||||||||||||||||||||| |||||| ||||||||||||||||||
Qy    377 GAACCTCAAAAGAGTCCCTGGTGTGAAGCAAGATCGCTAGAACACACCTTACCTGTAAAC 436

Db    649 agagagacactgaagaggaagcttagtgggagcagttgtgccagccctgttactagtcca 708
          |||||||||||||| ||||||| |||||||||| | |||| |||||||||||||| |||||
Qy    437 AGAGAGACACTGAAAAGGAAGGTTAGTGGGAACCGTTGCGCCAGCCCTGTTACTGGTCCA 496

Db    709 aacgcaaagagggatgctcacttctgcccggtctgcagcgattatgcatctgggtatcat 768
              ||||||||||||||||||||||| | |||||||||||||| ||||| || ||||| 
Qy    497 GGTTCAAAGAGGGATGCTCACTTCTGCGCTGTCTGCAGCGATTACGCATCGGGATATCAC 556

Db    769 tacggcgtttggtcatgtgaaggatgtaaggcctttttttaaaagaagcattcaaggacat 828
          || || || ||||| ||||||||||||||||||||||||||||||||||||||||||||
Qy    557 TATGGAGTCTGGTCGTGTGAAGGATGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACAT 616

Db    829 aatgattatatctgtccagccacgaatcagtgtaccatagacaagaaccggcgtaaaagc 888
          ||||||||||| |||||||| || |||||||||||| || || | || |||||||| || |||
Qy    617 AATGATTATATTTGTCCAGCTACAAATCAGTGTACAATCGATAAAAACCGGCGCAAGAGC 676

Db    889 tgccaggcctgccgacttcgcaagtgttatgaagtaggaatggtcaagtgtggatccagg 948
          |||||||||||||||||||| |||||||| ||||| |||||||| ||||||||| ||| ||
Qy    677 TGCCAGGCCTGCCGACTTCGGAAGTGTTACGAAGTGGGAATGGTGAAGTGTGGCTCCCGG 736

Db    949 agagaacggtgtgggtaccgtatagtgcggaggcagagaagttctagcgagcaggtacac 1008
          |||||  | |||||||||||  | |||||||| ||||||||||| |   ||||||| | |||
Qy    737 AGAGAGAGATGTGGGTACCGCCTTGTGCGGAGACAGAGAAGTGCCGACGAGCAGCTGCAC 796

Db   1009 tgcctgagcaaagccaagagaaacggtgggcatgcaccccgggtgaaggagctactgctg 1068
          ||     |||| ||||||||||  || || |||| || |||||| |||  |||||| ||||||
Qy    797 TGTGCCGGCAAGGCCAAGAGAAGTGGCGGCCACGCGCCCCGAGTGCGGGAGCTGCTGCTG 856

Db   1069 agcacttgagtccagagcaactggtgctcaccctcctggaagctgaaccagcccaatgtg 1128
            | | |||| || ||||| || |||||||||||||||||| ||||| ||  ||| |||||
Qy    857 GACGCCTGAGCCCCGAGCAGCTAGTGCTCACCCTCCTGGAGGCTGAGCCCCCCCCATGTG 916

Db   1129 ctggtgagccgtcccagcatgcccttcacggaggcctccatgatgatgtccctcactaag 1188
          ||| | |||| |||||   ||||||||| |||||||||||||||||||||||| || |||
Qy    917 CTGATCAGCCGCCCCAGTGCGCCCTTCACCGAGGCCTCCATGATGATGTCCCTGACCAAG 976

Db   1189 cttgcggacaaggagctggtgcacatgattggctgggccaagaaaatccctggctttgtg 1248
           | || |||||||||| |||| |||||||| |||||||||||||| || || |||||||||
Qy    977 TTGGCCGACAAGGAGTTGGTACACATGATCAGCTGGGCCAAGAAGATTCCCGGCTTTGTG 1036

Db   1249 gagctcagcctgtttgaccaagtccggctcttagaaagctgctggatggaggtgctaatg 1308
          |||||||||||||| |||||||| |||||||| || |||||| |||||||||||||| |||||
Qy   1037 GAGCTCAGCCTGTTCGACCAAGTGCGGCTCTTGGAGAGCTGTTGGATGGAGGTGTTAATG 1096

Db   1309 gtgggactgatgtggcgctcaattgaccaccccggcaagctcatcttcgctccagacctc 1368
           |||| |||||||||||||||||||||||||||||||||||||||||| ||||||||| || 
Qy   1097 ATGGGGCTGATGTGGCGCTCAATTGACCACCCCGGCAAGCTCATCTTTGCTCCAGATCTT 1156

Db   1369 gttctggacagggatgaggggaagtgcgtagaagggattctggaaatctttgacatactc 1428
          |||||||||||||||||||||||| |||||||||||||| |||||||||||||||||| |||
Qy   1157 GTTCTGGACAGGGATGAGGGGAAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTC 1216

Db   1429 ctggcgacgacgtcaaggttccgtgagttaaaactccagcacaaggagtatctctgtgtg 1488
          ||||| || || |||||||| || ||||||||||||| |||| || ||||||||||| 
Qy   1217 CTGGCAACTACTTCAAGGTTTCGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTC 1276

Db   1489 aaggccatgatcctcctcaactccagtatgtaccccttggcttctgcaaaccaggaggca 1548
          |||||||||||||| ||||| ||||||||||||||  |||   | || | |||||| || 
Qy   1277 AAGGCCATGATCCTGCTCAATTCCAGTATGTACCCTCTGGTCACAGCGACCCAGGATGCT 1336

Db   1549 gaaagtagccggaagctgacacacctactgaacgcggtgacagatgccctggtctgggtg 1608
          || || |||||||||||| | ||| | |||||||| |||||| |||||  ||||| ||||||
Qy   1337 GACAGCAGCCGGAAGCTGGCTCACTTGCTGAACGCCGTGACCGATGCTTTGGTTTGGGTG 1396

Db   1609 attgcgaagagtggtatctcctcccagcagcagtcagtccgactggccaacctcctgatg 1668
          ||||| ||||| || |||||||||||||||||| ||  | || ||||| ||||||||||||
Qy   1397 ATTGCCAAGAGCGGCATCTCCTCCCAGCAGCAATCCATGCGCCTGGCTAACCTCCTGATG 1456

Db   1669 cttctttctcacgtcaggcacatcagtaacaagggcatggaacatctgctcagcatgaag 1728
          || || || ||||||||||    |||||||||||||||||||||||||||||| |||||||
Qy   1457 CTCCTGTCCCACGTCAGGCATGCGAGTAACAAGGGCATGGAACATCTGCTCAACATGAAG 1516

Db   1729 tgcaaaaatgtggtcccggtgtatgacctgctgctggagatgctgaatgctcacgcactt 1788
          ||||||||||||||||| |||||||||||||||||||||||||||||||| ||||  |||
Qy   1517 TGCAAAAATGTGGTCCCAGTGTATGACCTGCTGCTGGAGATGCTGAATGCCCACGTGCTT 1576

Db   1789 cgagggtacaagtcctcaatctcggggtctgagtgcagctcaacagaggacagtaagaac 1848
          || |||| |||||||||| ||| ||||||| ||||||||| |  ||||||||||||| | |
Qy   1577 CGCGGGTGCAAGTCCTCCATCACGGGGTCCGAGTGCAGCCCGGCAGAGGACAGTAAAAGC 1636

Db   1849 aaagagagctcccagaacctacagtctcagtgatggccaggcctgaggcggac 1901
          |||||| ||||||||||||| |||||||||||||| | |  | ||||||| | ||
Qy   1637 AAAGAGGGCTCCCAGAACCCACAGTCTCAGTGACGCCTGGCCCTGAGGTGAAC 1689


RESULT    3
ID   T62844 standard; cDNA; 1458 BP.
AC   T62844;
DT   08-JUN-1997  (first entry)
DE   Mouse oestrogen receptor beta (ER-beta) cDNA.
KW   Orphan receptor; oestrogen receptor beta; ER-beta;
KW   nuclear receptor; prostate cancer; benign prostatic hyperplasia;
KW   ovary cancer; cardiovascular disease; osteoporosis;
KW   environment; pollutant; ss.
OS   Mus sp.
PN   WO9709348-A2.
PD   13-MAR-1997.
PF   09-SEP-1996; E03933.
PR   08-SEP-1995; GB-018272.
PR   15-MAR-1996; GB-005550.
PR   11-APR-1996; GB-007532.
PR   09-MAY-1996; GB-009576.
PA   (KARO-) KARO BIO AB.
PI   Enmark E,  Gustafsson J,  Kuiper GG;
DR   WPI; 97-192842/17.
DR   P-PSDB; W14725.
PT   New isolated oestrogen receptor beta - used to develop prods. for
PT   treating e.g. cancers, CNS diseases, osteoporosis or cardiovascular
PT   disease
PS   Claim 6; Fig 14B; 45pp; English.
```

```
CC   A cDNA clone (T62844) codes for a novel mouse oestrogen receptor-
CC   related nuclear receptor (W13725), designated ER-beta.  Rat, human
CC   and mouse ER-beta (W14723-25) can be used to isolate molecules for
CC   use in the treatment of cardiovascular diseases, central nervous
CC   system diseases, osteoporosis, prostate or ovarian cancer or benign
CC   prostatic hyperplasia and to test environmental chemicals for
CC   oestrogenic activity
SQ   Sequence   1458 BP;    365 A;    373 C;    403 G;    317 T;

  Query Match             53.2%;  Score 1010;  DB 29;  Length 1458;
  Best Local Similarity   84.6%;  Pred. No. 0.00e+00;
  Matches 1234;  Conservative    0;  Mismatches 224;  Indels    0;  Gaps    0;

Db          1 atggcattctacagtcctgctgtgatgaactacagtgttcccagcagcaccggtaacctg 60
              ||| ||||||| || |||||||||||||| |||||  |||||||||    |  |||| ||
Qy        212 ATGACATTCTATAGCCCTGCTGTGATGAATTACAGCATTCCCAGCAATGTCACTAACTTG 271

Db         61 gaaggtgggcctgttcgccagactgcaagcccaaatgtgctatggccaacttctggacac 120
              ||||||||||||| ||| |||||  |||||||||||||  | ||||||||  |||| |||
Qy        272 GAAGGTGGGCCTGGTCGGCAGACCACAAGCCCAAATGTCTTGTGGCCAACACCTGGGCAC 331

Db        121 ctctctcctttagccacccactgccaatcatcgcttctctatgcagaacctcaaaagagt 180
              || ||||||||||    |||  |||| | ||| | ||| ||||| |||||||||||||||
Qy        332 CTTTCTCCTTTAGTGGTCCATCGCCAGTTATCACATCTGTATGCGGAACCTCAAAAGAGT 391

Db        181 ccttggtgtgaagcaagatcactagaacacaccttgcctgtaaacagagagaccctgaag 240
              || ||||||||||||||||| |||||||||||||| ||||||||||||||||| ||||| 
Qy        392 CCCTGGTGTGAAGCAAGATCGCTAGAACACACCTTACCTGTAAACAGAGAGACACTGAAA 451

Db        241 aggaagcttggcgggagcggttgtgccagccctgttactagtccaagcaccaagagggat 300
              |||||| || | |||| | |||| ||||||||||||||| ||||| |  | |||||||||
Qy        452 AGGAAGGTTAGTGGGAACCGTTGCGCCAGCCCTGTTACTGGTCCAGGTTCAAAGAGGGAT 511

Db        301 gctcacttctgtgccgtctgcagtgattatgcatctgggtatcattacggtgtctggtcc 360
              ||||||||||| || |||||||| ||||| ||||| || ||||| || || |||||||| 
Qy        512 GCTCACTTCTGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCG 571

Db        361 tgtgaaggatgtaaggccttttttaaaagaagcattcaaggacataatgactatatctgt 420
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||| |||
Qy        572 TGTGAAGGATGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGT 631

Db        421 ccagccacgaatcagtgtacgatagacaagaaccggcgtaaaaactgccaggcctgccga 480
              ||||| || ||||||||||| || || || |||||||| || | ||||||||||||||||
Qy        632 CCAGCTACAAATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGA 691

Db        481 cttcgcaagtgttacgaagtaggaatggtcaagtgtggatccaggagagaaaggtgtggg 540
              ||||| |||||||||||||| |||||||| |||||||| ||| ||||||| || ||||||
Qy        692 CTTCGGAAGTGTTACGAAGTGGGAATGGTGAAGTGTGGCTCCGGGAGAGAGAGATGTGGG 751

Db        541 taccgaatagtacgaagacagagaagtgccagcgagcaggtgcattgcctgaacaaagcc 600
              |||||   | || || |||||||||||||||    ||||||| |||| ||     ||| |||
Qy        752 TACCGCCTTGTGCGGAGACAGAGAAGTGCCGACGAGCAGCTGCACTGCCCGGGCAAGGCC 811

Db        601 aagagaaccagtgggcacacaccccgggtgaaggagctactgctgaactctctgagtccc 660
              |||||||     | || ||| | ||||| |||    |||||| |||||| || | ||||| |||
Qy        812 AAGAGAAGTGGCGGCCACGGTGCCCCGAGTGCGGGAGCTCCTGCTGCACCCCTTGAGCCC 871

Db        661 gagcagctggtgctcaccctgctggaggctgagccacccaatgtgctagtgagtcgtccc 720
              |||||||| |||||||||| ||||| |||||||| ||| ||||||  | || || |||
Qy        872 GAGCAGCTATTGCTCACCCTCCTGGAGGCTGAGCCGCCCCATGTGCTGATCAGCCGCCCC 931

Db        721 agcatgcccttcaccgaggcctccatgatgatgtcccttacgaagctggctgacaaggaa 780
              ||    |||||||||||||||||||||||||||||||| || ||| |||| ||||||||
Qy        932 AGTGCGCCCTTCACCGAGGCCTCCATGATGATGTCCCTGACCAAGTTGGCGGACAAGGAG 991

Db        781 ctggtgcacatgattggctgggccaagaaaatccctggctttgtggagctcagcctgttg 840
               |||| ||||||||  |||||||||||||| || || |||||||||||||||||||||||
Qy        992 TTGGTACACATGATCAGCTGGGCCAAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTC 1051

Db        841 gaccaagtccgcctcttggaaagctgctggatggaggtgctgatggtggggctgatgtgg 900
              |||||||| || ||||||||| ||||| ||||||||||||| | ||| ||||||||||||||
Qy       1052 GACCAAGTGCGCCTCTTCGAGAGCTGTTGGATGGAGGTGCTCAATGATGGGCTGATGTGG 1111

Db        901 ggctccatcgaccagccccgcaagctcatctttgctccagacctcgttctggacagggat 960
              ||||| || |||||||||:|||||||||||||||||||||||||| || ||||||||||||
Qy       1112 GGCTCAATTGACCAGCCCCGCAAGCTCATCTTTGCTCCAGACCTCATTTTGGACAGGGAT 1171

Db        961 gaggggaagtgcgtggaaggaattctggaaatctttgacatgctcctggcaacgtcctca 1020
              |||||||| ||||| ||||| |||||||||||||||||||||||||||||||| || ||   ||
Qy       1172 GAGGGGAAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTACTTCA 1231

Db       1021 cggttcggtgagttaaaactgcagcacaaagaatatctgtgtgtgaaggccatgattctc 1080
               |||| || ||||||||||| || |||||||||||||| ||||| |||||||||||| ||
Qy       1232 AGCTTTCGAGAGTTAAAACTCCAACATAAAGAATATCTCTGTGTCAAGGCCATCATCCTC 1291

Db       1081 ctcaactccagtatgtaccacttggctaccgcaaccaggaagcagagagtagcccggaag 1140
              ||||| |||||||||||||    |||   || || | |||||| || || || ||||||||
Qy       1292 CTCAATTCCAGTATGTACCCTCTGGTCACAGCGACCCAGCATCCTGACACCACCCCGAAG 1351

Db       1141 ctgacacacctattgaacgcagtgacagatgccctggtctgggtgatttcaaagagcacc 1200
              ||| | |||| |  ||||||| ||||| |||||  |||| ||||||||| | |||||   |
Qy       1352 CTGGCTCACTTGCTGAACGCCGTGACCGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGC 1411

Db       1201 atctcttcccagcagcagtcagtccgtctggccaacctcctgatgcttctttctcatgtc 1260
              ||||| ||||||||||| ||  | || ||||| |||||||||||||||| || || || |||
Qy       1412 ATCTCCTCCCAGCAGCAATCCATGCGCCTGGCTAACCTCCTGATGCTCCTGTCCCACGTC 1471

Db       1261 aggccatcaggtaacaagggcatggaacatctgctcagcatgaagtgcaaaaatgtggtc 1320
              |||||      ||||||||||||||||||||||||||| ||||||||||||||||||||||
Qy       1472 AGGCATGCGAGTAACAAGGGCATGGAACATCTGCTCAACATGAAGTGCAAAAATGTGGTC 1531

Db       1321 ccggtgtacgacctgctgctggagatgctgaatgctcacacgcttcgaggtacaagtcc 1380
              || ||||| |||||||||||||||||||||||||| |||   |||||| |||| ||||||
Qy       1532 CCAGTGTATGACCTGCTGCTGGAGATGCTGAATGCCCACGTGCTTCGCGGGTGCAAGTCC 1591

Db       1381 tcaatctcggggtctggtgctgctcgacaggagacagtaaagagcaaaaagggctcccag 1440
              || ||| ||||||| | |||| || || ||||||||||||| |||||||||||||||||||
Qy       1592 TCCATCACGGGGTCCGAGTGCAGCCCGGCAGAGGACAGTAAAAGCAAAGAGGGCTCCCAG 1651

Db       1441 aacctccagtctcagtga 1458
              ||||  ||||||||||||
Qy       1652 AACCCACAGTCTCAGTGA 1669


RESULT    4
ID   T06873 standard; DNA; 4963 BP.
AC   T06873;
DT   13-MAR-1996  (first entry)
DE   PhCMV*-1 promoter and human oestrogen receptor.
KW   Transactivator, tTA, tet operator, tetO, gene expression;
KW   tetracycline-responsive promoter; PhCMV*-1; human cytomegalovirus;
KW   transcription activation; eukaryotic cell; oestrogen receptor;
KW   ds; ss.
OS   Chimeric synthetic;
OS   Chimeric Homo sapiens.
PN   US5464758-A.
PD   07-NOV-1995.
PF   14-JUN-1993; 076726.
PR   14-JUN-1993; US-076726.
PA   (BUJA/) BUJARD H.
PA   (GOSS/) GOSSEN M.
PI   Bujard H,  Gossen M;
DR   WPI; 95-392512/50.
PT   Polynucleotide encoding transactivator fusion protein contg. tet
PT   repressor  - used to control expression of gene regulated by minimal
PT   promoter linked to tet operon, and vectors and cells where gene
PT   expression is regulated by tetracycline
PS   Disclosure; Fig 10A-B; 37pp; English.
CC   cDNA coding for the human oestrogen receptor was placed under
CC   control of the PhCMV*-1 promoter (see T06869), resulting in
CC   tetracycline-regulated expression vector pUHDpgr-3 contg. the
CC   sequence given in T06873.  The construct was transfected into
CC   eukaryotic cells expressing the tTA transactivator (T06867) to
CC   allow tetracycline-regulated expression of the receptor.
SQ   Sequence   4963 BP;   1259 A;   1300 C;   1278 G;   1126 T;
```

```
  Query Match             10.4%;  Score 197;  DB 17;  Length 4963;
  Best Local Similarity   66.7%;  Pred. No. 6.70e-125;
  Matches  424;  Conservative    0;  Mismatches 209;  Indels    3;  Gaps    2;

Db   1471 cccttcagtgaagcttcgatgatgggcttactgaccaacctggcagacagggagctggtt 1530
          |||||||  || || || ||||||   |  ||||||||  |||| |||| |||| ||||
Qy    938 CCCTTCACCGAGGCCTCCATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTA 997

Db   1531 cacatgatcaactgggcgaagagggtgccaggctttgtggatttgaccctccatgatcag 1590
          |||||||||| |||||| |||| | | || |||||||||||  | | |||    || ||
Qy    998 CACATGATCAGCTGGGCCAAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAA 1057

Db   1591 gtccaccttctagaatgtgcctggctagagatcctgatgattggtctcgtctggcgctcc 1650
          || |  ||  | ||  |    ||| | ||| |  | ||||| || ||  | ||||||||
Qy   1058 GTGCGGCTCTTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCA 1117

Db   1651 atggagcacccagtgaagctactgtttgctcctaacttgctcttggacaggaaccaggga 1710
          || || ||||| |  |||||  | ||||||||  |  |  |  |||||||| |  ||||
Qy   1118 ATTGACCACCCCGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGG 1177

Db   1711 aaatgtgtagagggcatggtggagatcttcgacatgctgctggctacatcatctcggttc 1770
          ||||| ||||| || ||  || | ||||| |||||||| ||||| ||  | ||  ||||
Qy   1178 AAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTT 1237

Db   1771 cgcatgatgaatctgcagggagaggagtttgtgtgcctcaaatctattattttgcttaat 1830
          ||   | | || || ||     | || | | | ||  ||||  | || ||  |||| |||
Qy   1238 CGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAAT 1297

Db   1831 tctggagtgtacacatttctgtccagcaccctgaagtctctggaagagaaggaccatatc 1890
          ||  |  ||||| |  |  |  |    |||| | |  ||    |  ||  |||   |  |
Qy   1298 TCCAGTATGTACCCTCTGGTCACAGCGACCCAGGATGCTGAC-AGCAGCCGGAAGCTGGC 1356

Db   1891 caccgagtcctggacaagatcacagacactttgatccacctgatggccaaggcaggcctg 1950
             |  |  ||  ||    | || ||  ||||| |     |||| ||||||   ||| |
Qy   1357 TCACTTG--CTCAACCCCGTGACCGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATC 1414

Db   1951 accctgcagcagcagcaccagcggctggcccagctcctcctcatcctctcccacatcagg 2010
           ||   ||||||||   |  ||| |||||  | |||||  |  |||| |||||| |||||
Qy   1415 TCCTCCCAGCAGCAATCCATGCGCCTGGCTAACCTCCTGATGCTCCTGTCCCACGTCAGG 1474

Db   2011 cacatgagtaacaaaggcatggagcatctgtacagcatgaagtgcaagaacgtggtgccc 2070
          ||   ||||||||| |||||||| ||||||  || |||||||||||| || ||||| ||
Qy   1475 CATGCGAGTAACAAGGGCATGGAACATCTGCTCAACATGAAGTGCAAAAATGTGGTCCCA 1534

Db   2071 ctctatgacctgctgctggagatgctggacgcccac 2106
           | |||||||||||||||||||||||| | ||||||
Qy   1535 GTGTATGACCTGCTGCTGGAGATGCTGAATGCCCAC 1570


RESULT    5
ID   Q76270 standard; DNA; 4963 BP.
AC   Q76270;
DT   17-JUL-1995  (first entry)
DE   PhCMV*-1-controlled progesterone receptor
KW   tTA; transactivator; tetracycline-controllable transactivator;
KW   conditional inactivation; homologous recombination; gene expression;
KW   gene regulation; gene therapy; tetracycline-resistance; tetR;
KW   transgenic animal; PhCMV*-1; promoter; tetO; CMV;
KW   progesterone receptor; ds.
OS   Human cytomegalovirus Towne.
PN   WO9429442-A.
PD   22-DEC-1994.
PF   14-JUN-1994; U06734.
PR   14-JUN-1993; US-076327.
PA   (BADI ) BASF AG.
PI   Bujard H,  Gossen M,  Salfeld JG,  Voss JW;
DR   WPI; 95-036472/05.
PT   Regulatory systems using tetracycline-controllable transactivator
PT   (tTA) - useful for conditional inactivation or modulation of
PT   gene expression in a host cell or animal
PS   Disclosure; Page 54-59; 102pp; English.
CC   The minimal promoter PhCMV*-1 is based on the human CMV
CC   intermediate-early promoter (with enhancer region removed), fused
CC   to a tetracycline operator, tetO.  The promoter is used to
CC   express tTA transactivator in host cells.  In clone pUHD 5GR4
CC   (given in Q76269), cDNA for rabbit progesterone receptor is under
CC   the control of PhCMV*-1.
SQ   Sequence   4963 BP;    1259 A;    1298 C;    1280 G;    1126 T;

  Query Match             10.4%;  Score 197;  DB 13;  Length 4963;
  Best Local Similarity   66.7%;  Pred. No. 6.70e-125;
  Matches  424;  Conservative    0;  Mismatches 209;  Indels    3;  Gaps    2;

Db   1471 cccttcagtgaagcttcgatgatgggcttactgaccaacctggcagacagggagctggtt 1530
          |||||||  || || || ||||||   |  ||||||||  |||| |||| |||| ||||
Qy    938 CCCTTCACCGAGGCCTCCATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTA 997

Db   1531 cacatgatcaactgggcgaagagggtgccaggctttgtggatttgaccctccatgatcag 1590
          |||||||||| |||||| |||| | | || |||||||||||  | | |||    || ||
Qy    998 CACATGATCAGCTGGGCCAAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAA 1057

Db   1591 gtccaccttctagaatgtgcctggctagagatcctgatgattggtctcgtctggcgctcc 1650
          || |  ||  | ||  |    ||| | ||| |  | ||||| || ||  | ||||||||
Qy   1058 GTGCGGCTCTTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCA 1117

Db   1651 atggagcacccagtgaagctactgtttgctcctaacttgctcttggacaggaaccaggga 1710
          || || ||||| |  |||||  | ||||||||  |  |  |  |||||||| |  ||||
Qy   1118 ATTGACCACCCCGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGG 1177

Db   1711 aaatgtgtagagggcatggtggagatcttcgacatgctgctggctacatcatctcggttc 1770
          ||||| ||||| || ||  || | ||||| |||||||| ||||| ||  | ||  ||||
Qy   1178 AAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTT 1237

Db   1771 cgcatgatgaatctgcagggagaggagtttgtgtgcctcaaatctattattttgcttaat 1830
          ||   | | || || ||     | || | | | ||  ||||  | || ||  |||| |||
Qy   1238 CGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAAT 1297

Db   1831 tctggagtgtacacatttctgtccagcaccctgaagtctctggaagagaaggaccatatc 1890
          ||  |  ||||| |  |  |  |    |||| | |  ||    |  ||  |||   |  |
Qy   1298 TCCAGTATGTACCCTCTGGTCACAGCGACCCAGGATGCTGAC-AGCAGCCGGAAGCTGGC 1356

Db   1891 caccgagtcctggacaagatcacagacactttgatccacctgatggccaaggcaggcctg 1950
             |  |  ||  ||    | || ||  ||||| |     |||| ||||||   ||| |
Qy   1357 TCACTTG--CTCAACCCCGTGACCGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATC 1414

Db   1951 accctgcagcagcagcaccagcggctggcccagctcctcctcatcctctcccacatcagg 2010
           ||   ||||||||   |  ||| |||||  | |||||  |  |||| |||||| |||||
Qy   1415 TCCTCCCAGCAGCAATCCATGCGCCTGGCTAACCTCCTGATGCTCCTGTCCCACGTCAGG 1474

Db   2011 cacatgagtaacaaaggcatggagcatctgtacagcatgaagtgcaagaacgtggtgccc 2070
          ||   ||||||||| |||||||| ||||||  || |||||||||||| || ||||| ||
Qy   1475 CATGCGAGTAACAAGGGCATGGAACATCTGCTCAACATGAAGTGCAAAAATGTGGTCCCA 1534

Db   2071 ctctatgacctgctgctggagatgctggacgcccac 2106
           | |||||||||||||||||||||||| | ||||||
Qy   1535 GTGTATGACCTGCTGCTGGAGATGCTGAATGCCCAC 1570


RESULT    6
ID   N70880 standard; cDNA; 2092 BP.
AC   N70880;
DT   24-FEB-1991  (first entry)
DE   cDNA encoding human oestrogen receptor protein.
KW   Human oestrogen receptor protein; ss cDNA; expression system; assay;
KW   steroid receptor protein;
OS   Homo sapiens.
FH   Key                 Location/Qualifiers
FT   CDS                 293..1280
FT   /*tag=  a
PN   WO8705049-A.
PD   27-AUG-1987.
PF   18-FEB-1987; U00341.
PR   20-FEB-1986; US-833929.
PA   (CALB-) California Biotechnology Inc.
PI   Shine J;
DR   WPI; 87-250013/35.
DR   P-PSDB; P70543.
PT   Expression system for vertebrate steroid receptor protein - comprising
PT   DNA sequence encoding the protein linked to control sequences in
PT   eucaryotic hosts.
PS   Disclosure; Fig [illegible]; [illegible]pp; English
CC   The cDNA can be used in an expression system to express human
CC   oestrogen receptor protein.  The coding sequence is operably linked to
CC   control sequences compatible with eukaryotic host cells.  This method
CC   allows expression under conditions which favour appropriate post-
CC   translational processing.  It produces large amts. of purified protein
CC   useful in the design of agonist and antagonist cpds., for the study of
CC   the mechanism of action of the steroid binding proteins in general,
CC   and for use in diagnostic assays for the proteins or antibodies to
CC   them.  These assays are important in, eg the diagnosis of tumour
CC   sensitivities to steroid metabolism.  Suitable host cells are VERO,
CC   HeLa and CHO cells.
SQ   Sequence   2092 BP;    473 A;    607 C;    591 G;    421 T;

  Query Match             10.3%;  Score 195;  DB 2;  Length 2092;
  Best Local Similarity  66.5%;  Pred. No. 2.29e-123;
  Matches  422;  Conservative    0;  Mismatches 213;  Indels    3;  Gaps    2;

Db     1298 cccttcagtgaagcttcgatgatgggcttactgaccaacctggcagacagggagctggtt 1357
            |||||||  || || || ||||||   |  ||||||||  |||| |||| |||| ||||
Qy      938 CCCTTCACCGAGGCCTCCATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTA 997

Db     1358 cacatgatcaactgggcgaagagggtgccaggctttgtggatttgaccctccatgatcag 1417
            |||||||||| |||||| ||||| | | || ||||||||||||  | | |||    || ||
Qy      998 CACATGATCAGCTGGGCCAAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAA 1057

Db     1418 gtccaccttctagaatgtgcctggctagagatcctgatgattggtctcgtctggcgctcc 1477
            || |  ||  | ||  |    ||| | ||| |  | ||||| || ||  | ||||||||
Qy     1058 GTGCGGCTCTTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCA 1117

Db     1478 atggagcacccagtgaagctactgtttgctcctaacttgctcttggacaggaaccaggga 1537
            || || |||||| |  |||||  | ||||||||  |  |  ||  ||||||||| |  ||||
Qy     1118 ATTGACCACCCCGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGG 1177

Db     1538 aaatgtgtagagggcatggtggagatcttcgacatgctgctggctacatcatctcggttc 1597
            ||||| ||||| || ||  |||| ||||| |||||||| ||||| ||  | ||  ||||
Qy     1178 AAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTT 1237

Db     1598 cgcatgatgaatctgcagggagaggagtttgtgtgcctcaaatctattatttgcttaat 1657
            ||   | | || || ||     | || | | | ||  |||| | || ||  |||| |||
Qy     1238 CGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAAT 1297

Db     1658 tctggagtgtacacatttctgtccagcaccctgaagtctctggaagagaaggaccatatc 1717
            ||  |  ||||| |  |  |  |    |||| | |  ||    |  ||  |||   |  |
Qy     1298 TCCAGTATGTACCCTCTGGTCACAGCGACCCAGGATGCTGAC-AGCAGCCGGAAGCTGGC 1356

Db     1718 caccaagtcctggacaagatcacagacactttgatccacctgatggccaaggcaggcctg 1777
               |  |  ||| ||    | || ||  ||||| |     |||| ||||||   ||| |
Qy     1357 TCACTTG--CTGAACGCCGTGACCGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATC 1414

Db     1778 accctgcagcagcagcaccagcggctggcccagctcctcctcatcctctcccacatcagg 1837
             ||   ||||||||   |  ||||||||| | |||||  |  |||| ||||||| |||||
Qy     1415 TCCTCCCAGCAGCAATCCATGCGGCTGGCTAACCTCCTGATGCTCCTGTCCCACGTCAGG 1474

Db     1838 cacatgagtaacaaaggcatggagcatctgtacagcatgaagtgcaagaacgtggtgccc 1897
            ||   |||||||||| ||||||||| ||||||  || |||||||||||| || ||||| ||
Qy     1475 CATGCGAGTAACAAGGGCATGGAACATCTGCTCAACATGAAGTGCAAAAATGTGGTCCCA 1534

Db     1898 ctctatgacctgctgctggagatgctggacgcccac 1933
             | ||||||||||||||||||||||||| | ||||||
Qy     1535 GTGTATGACCTGCTGCTGGAGATGCTGAATGCCCAC 1570


RESULT     7
ID   Q85976 standard; cDNA; 204 BP.
AC   Q85976;
DT   31-OCT-1995  (first entry)
DE   Human estrogen receptor cDNA.
KW   nuclear receptor protein; osteosarcoma; EAR-1r; generation; growth;
KW   differentiation; human; erb-A related gene; estrogen receptor; ds.
OS   Homo sapiens.
PN   J07033798-A.
PD   03-FEB-1995.
FF   20-JUL-1993; 200333.
PP   20-JUL-1993; JP-200333.
PA   (MITU ) MITSUBISHI KASEI CORP.
DR   WPI; 95-110650/15.
DR   P-PSDB; R71566.
PT   Nuclear receptor protein derived from osteoblastic osteosarcoma
PT   - encoded by the EAR-1r gene, for investigation of generation of
PT   fat soluble active substances
PS   Disclosure; Page 6; 8pp; Japanese.
CC   This cDNA encodes an estrogen receptor (a member of a receptor
CC   superfamily). It is disclosed in the patent specification. The
CC   invention concerns an EAR-1r gene (human erb-A related gene)
CC   encoding a nuclear receptor protein (R71565) derived from human
CC   osteosarcoma (HOS) cells. The expression of EAR-1r is most abundant
CC   in the brain and was detectedd slightly in heart and placenta. The
CC   receptor protein may be used for the investigation of generation,
CC   differentiation and growth of fat soluble physiological substances.
SQ   Sequence   204 BP;    59 A;    46 C;    55 G;    44 T;

  Query Match              6.7%;  Score 127;  DB 14;  Length 204;
  Best Local Similarity  81.3%;  Pred. No. 8.13e-72;
  Matches  165;  Conservative    0;  Mismatches   38;  Indels    0;  Gaps    0;

Db        1 tgtgcagtgtgcaatgactatgcttcaggctaccattatggagtctggtcctgtgagggc 60
            || || || ||||  || || || || || || || |||||||||||||| ||||| ||
Qy      521 TGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGA 580

Db       61 tgcaaggccttcttcaagagaagtattcaaggacataacgactatatgtgtccagccacc 120
            || |||||||| || || ||||| |||||||||||||| || ||||| |||||| ||
Qy      581 TGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACA 640

Db      121 aaccagtgcaccattgataaaaacaggaggaagagctgccaggcctgccggctccgcaaa 180
            || ||||| || || ||||||||| || | |||||||||||||||||||| || || ||
Qy      641 AATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAG 700

Db      181 tgctacgaagtgggaatgatgaa 203
            || ||||||||||||||| ||||
Qy      701 TGTTACGAAGTGGGAATGGTGAA 723


RESULT     8
ID   Q10572 standard; DNA; 1047 BP.
AC   Q10572;
DT   09-APR-1991  (first entry)
DE   Human Natriuretic Peptide Receptor B.
KW   NPRB; ANP; BNP; CNP; kidney failure; heart failure; protein kinase;
KW   hyperaldosteronism; glaucoma; guanyl cyclase.
OS   Homo sapiens.
FH   Key                Location/Qualifiers
FT   Peptide            1..22
FT   /label= signal sequence
FT   Protein            12
FT   /label= mature NPBR
FT   Domain             23..455
FT   /label= extracellular domain
FT   /note= "binds natriuretic peptides A,B and C]"
FT   Domain             456..456
FT   /label= transmembrane domain
FT   Domain             479..1047
FT   /label= cytoplasmic domain
FT   /note= "GC and protien kinase activity"
FT   Modified -site  24..26
FT   /label= N-glycos_site
FT   Modified -site  35..37
FT   /label= N-glycos_site
FT   Modified -site  161..163
FT   /label= N-glycos_site
FT   Modified -site  195..197
FT   /label= N-glycos_site
FT   Modified -site  244..246
FT   /label= N-glycos_site
FT   Modified -site  277..279
FT   /label= N-glycos_site
FT   Modified -site  349..351
FT   /label= N-glycos_site
FT   Modified -site  600..602
FT   /label= N-glycos_site
PN   WO9100292-A.
PD   10-JAN-1991.
PF   22-JUN-1990; U03586.
PR   23-JUN-1989; US-370673.
PA   (GETH ) GENENTECH INC.
PI   Chang M  Goeddel D  Lowe D;
DR   WPI; 91-036711/05.
DR   N-PSDB; Q10324.
PT   Natriuretic protein receptor B - for diagnosis and treatment of
PT   kidney failure, heart failure, hyperaldosteronism, glaucoma etc
PS   Claim 3; Fig 1; 49pp; English.
CC   The sequence was derived from the DNA encoding natriuretic peptide
CC   receptor B, NPPB, having guanyl cyclase (GC) activity and protein
CC   kinase activity.  The DNA can be inserted into expression vectors
CC   for the prodn. of the protein, opt. after being mutated to produce
CC   NPRB analogues  The protein has a mol wt  of 115 kD (calculated Mr=
CC   114,952).   The protein (or variants) can be used in treatment of
CC   natriuretic peptide disorders, and also to isolate peptides using
CC   affinity chromatography.  Antibodies with affinity for NPRB can
CC   also be prepd.
SQ   Sequence   1047 BP;    87 A;    15 C;    83 G;    51 T;

  Query Match              4.8%;  Score 91;  DB 2;  Length 1047;
  Best Local Similarity   8.9%;  Pred. No. 2.71e-45;
  Matches    83;  Conservative  269;  Mismatches 572;  Indels  10;  Gaps  10;

Db     108 varnashwrnnnntagavasgnsakndhyrtnvrtgnsankngnnvvtnhghnnwtaraa 167
            :|: |            || | |  | .     |         .    ..  :  :  :
Qy     433 AAACAGAGAGACACTGAAAAGGAAGGTTAGTGGGAACCGTTGCGCCAGCCCTGTTACTGG 492

Db     168 nnyndart-ddrnhyntnngvnnanngsnnsvnhnvyarnnggnnathnnrangrnvyn 226
              : :  | ...  .        |.  |    .  .. .  .  .  |         .  .  |. :.
Qy     493 TCCAGGTTCAAAGAGGGATGCTCACTTCTGCCCTGTCTGCAGCGATTACGCATCGGGATA 552

Db     227 cgnnnmnhnnnnnanrnnntn-gdyvnnyndvngnsnragntratgrnwndnrtrnnana 285
                :        .   |  |.  .   :  . . |  .      | |  .  :  :  .|   |
Qy     553 TCACTATGGAGTCTGGTCGTGTGAAGGATGTAAGGCCTTTTTTAAAACAAGCATTCAAGG 612

Db     286 nrnanntvnvntyrnnnnnnynnnnnrnnnrarndngvnngusnmnnnagcnydgnnnya 345
                |:    ::    :                 : :  :    : :   |    ::|    |
Qy     613 ACATAATGATTATATTTGTCCAGCTACAAATCAGTGTACAATCGATAAAAACCGGCGCAA 672

Db     346 nvnnntnnnggtrndgnrnvnkmngrryhgvtgnvvmdknndrntdnvnwamgdndsgdn 405
             .      ||     .        . |.   | |             .   . |:    .|       .  .
Qy     673 GAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTACGAAGTGGGAATGGTGAAGTGTGGCTC 732

Db     406 nnaahysganknnwwtgrnnnwvkgannsdnnncandnddnscdktnnstnanvangtgn 465
                :  ||     :  ||:       ..      ..       |  .  . : .  . ::|   .   |  .|    |
Qy     733 CCGGAGAGAGAGATGTGGTACCGCCTTGTGGGAGATAGAGAAGTGCC-GACGAGCAGC 791

Db     466 tnnmngvssnnnnrknmnnknnasmnwrnrwnnnnngnsnryhkgagsrntnsnrgssyg 525
            |   :   :  :       :    :        |::  ..  ..                ..      . .  |  .  .|  ..
Qy     792 TGCACTGTGCCG-GCAAGGCCAAGAGAA-GTGGCGGCCACGCGCCCCGAGTGCGGGAGCT 849

Db     526 snmtahgkynnnantghnkgnvvankhvnkkrnnntrnvnnnnkhmrdvnnnhntrnnga 585
            :     :|:         ||.          :                 |         .
Qy     850 GCTGCTGGACGCCCTGAGCCCCGAGCAGCTAGTGCTCACCCTCCTGGAGGCTGAGCCGCC 909

Db     586 cndnnnncnvtnycnrgsnndnnnndsnnndwmnrysnnndnvkgmannhnsnnsshgsn 645
            | :       | :  :      .        ..        .     .               |  :  :   .  :
Qy     910 CCATGTGCTGATCAGCCGCCCCAGTGCGCCCTTCACCGAGGCCTCCATGATGATGTCCCT 969

Db     646 kssnnvvdsrnvnkntdygnasn-rstanndnnanyakknntannnnsgnnnnttgmna 704
            :  .         :  .  .  .  : |  |.          ||        |  |                 .   |
Qy     970 GACCAAGTTGGCCGACAAGGAGTTGGTACACATGATCAGCTGGGCCAAGAAGATTCCCGG 1029
```

```
Db     705 adyysngnnnnnanrsgnnynngndnsnknnvnkvrngnniynnisndrtn--nnnvnn 764
              .  .. |         |  ..    .  |        .  :  ::       :      :   :  :  |       :
Qy    1030 CTTTGTGGAGCTCA-GCCTGTTCGACCAAGTGCGGCTCTTGGAGAGCTGTTGGATGGAGG 1088

Db     765 mnrrwandnanrndngnnkgnnrrnnknggtsnndnnnnnrmnngannnnknvnnrtnayn 824
                :|  :    :  :      :|   :     :      |              :  :|                 |  :
Qy    1089 TGTTAATGATGGGGCTGATGTGGCGCTCAATTGACCACCCCGGCAAGCTCATCTTTGCTC 1148

Db     825 nnkrkanannynnnnhsvannnkrgntvnanandsvtnynsdnvgntans-anstnmnvv 883
              . . :       .       : . :        : .        .  | |  . . . |      :  :    |    :   |    |       :
Qy    1149 CAGATCTTGTTCTGGACAGGGATGAGGGGAAATGCGTAGAAGGAATTCTGGAAATCTTTG 1208

Db     884 tnnndnytcndanndnndvykvntngdaymvvsgnngrngnrhannnarmananndavss 943
               .  .  |  .    .    . . . . .       |. . . . . .      .         :           : :|  |   :  :  :
Qy    1209 ACATGCTCCTGGCAACTACTTCAAGGTTTCGAGAGTTAAAACTCCAACACAAAGAATATC 1268

Db     944 nrnrhrnhdnnrn-rnqvhtgnvcagvvgnkmnrycnngdtvntasrmnsnqnankndhvs 1002
                ::: ::    :    :  |::      .|    ..      .  ::         :  :  |    :     :          :  :::
Qy    1269 TCTGTGTCAAGGCCATGATCCTGCTCAATTCCAGTATGTACCCTCTGGTCACAGCCACCC 1328

Db    1003 -sttkdandnngcnnnnnrgdvnmkgkgkmrtyw 1035
              .   ..   .           .|. .  ..|.  . . .  ::
Qy    1329 AGGATGCTGACAGCAGCCGGAAGCTGGCTCACTT 1362


RESULT     9
ID   Q10572 standard, DNA, 1047 BP.
AC   Q10572;
DT   09-APR-1991  (first entry)
DE   Human Natriuretic Peptide Receptor B.
KW   NPRB, ANF, BNP, CNP; kidney failure, heart failure, protein kinase;
KW   hyperaldosteronism; glaucoma; guanyl cyclase.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   Peptide         1..22
FT   /label= signal sequence
FT   Protein         12
FT   /label= mature NPBR
FT   Domain          23..455
FT   /label= extracellular domain
FT   /note= "binds natriuretic peptides A,B and C]"
FT   Domain          456..456
FT   /label= transmembrane domain
FT   Domain          479..1047
FT   /label= cytoplasmic domain
FT   /note= "GC and protien kinase activity"
FT   Modified -site  24..26
FT   /label= N glycos_site
FT   Modified -site  35..37
FT   /label= N-glycos_site
FT   Modified -site  161..163
FT   /label= N-glycos_site
FT   Modified -site  195..197
FT   /label= N-glycos_site
FT   Modified -site  244..246
FT   /label= N-glycos_site
FT   Modified -site  277..279
FT   /label= N-glycos_site
FT   Modified -site  349..351
FT   /label= N-glycos_site
FT   Modified -site  600  602
FT   /label= N-glycos_site
PN   WO9100292-A.
PD   10-JAN-1991.
PF   22-JUN-1990; U03586.
PR   23-JUN-1989; US-370673.
PA   (GETH ) GENENTECH INC.
PI   Chang M, Goeddel D, Lowe D;
DR   WPI; 91-036711/05.
DR   N-PSDB, Q10324.
PT   Natriuretic protein receptor B - for diagnosis and treatment of
PT   kidney failure, heart failure, hyperaldosteronism, glaucoma etc.
```

```
PS   Claim 3; Fig 1; 49pp; English.
CC   The sequence was derived from the DNA encoding natriuretic peptide
CC   receptor B, NPRB, having guanyl cyclase (GC) activity and protein
CC   kinase activity.  The DNA can be inserted into expression vectors
CC   for the prodn. of the protein, opt. after being mutated to produce
CC   NPRB analogues. The protein has a mol wt. of 115 kD (calculated Mr=
CC   114,952).  The protein (or variants) can be used in treatment of
CC   natriuretic peptide disorders, and also to isolate peptides using
CC   affinity chromatography.  Antibodies with affinity for NPRB can
CC   also be prepd.
SQ   Sequence   1047 BP;    87 A;    15 C;    83 G;    51 T;

  Query Match              4.0%,  Score 75,  DB 2,  Length 1047,
  Best Local Similarity   8.7%;  Pred. No. 7.88e-34;
  Matches    46;  Conservative  161;  Mismatches 316;  Indels    6;  Gaps    6;

Db    442 dnddnsedktnnstnanvangtqntnnmngvssnnnnrknmnnknnasmnwrnrwnnnnn 501
          : :: :|: |   :| | :    ||     : |:::     :: .        .  .  . ..
Cp    553 ATATCCGATGCGTAATCGC-IGCAGACAGGCAGAAGTGAGCATCCCTCTTTGAACCTC 495

Db    502 qnsnr ykkqaqsrntnsnrqssyqsnmtahgkyannantqhnkgnvvankhvnkkrnnn 560
          | : : ::   ||:: | : :     :|:       .   .        | . .  ..  ... .
Cp    494 GACCAGTAACAGGGCTGGCGCAACGGTTCCCACTAACCTTCCTTTTCAGTGTCTCTCTGT 435

Db    561 trnvnnnnkhmrdvnnnhntrnnqaendnnnncnvtnycnrqsnndnnnndsnnndwmnr 620
          |  :      : :::    : |:  ||  :        :   :| :|:  :     ::    ::: :
Cp    434 TTACAGGTAAGGTGTGTTCTAGCGATCTTGCTTTACACCAGCCACCCCCCCAGCTTCCC 375

Db    621 ynnndnvkqmnnnhnnnnsshqsnkssnc vvdsrnvnkntdyqrasnrstannddnn 678
          :     : ::|    : :  :: : :  :: | ::::: :     ::|  : ::|    ::
Cp    374 CATACAGATGTGATAACTGGCGATGGACCACTAAAGGAGAAGGTGCCCAGGTGTTGCCC 315

Db    679 anyakknntannnnsqnnnnttgmnaadvysnqnnnnnnanrsgnnynngndnsnknnvn 738
          |  |    |     :       |      :::: |          :           :      :
Cp    314 ACAACACATTTGGGCTTGTGGTCTGCCGACCAGGCCCACCTTCCAACTTAGTGACATTGC 255

Db    739 kvrnqnrnynrnsndrtnnnnnnvnnmnrcwandnanrndnqnnkqnnrrnnknqgtsnn 798
          :::     :  :   :         :      :|:|  :  |    :     :    ::   :   |:
Cp    254 TGGGAATGCTGTAATTCATCACAGCAGGCTATAGAATGTCATGCTGCATATTCATGCT 195

Db    799 dnnnnrmnnyannnnknvnnrtnaynnnkrkanannynnnnhsvannnkrqntvnanand 858
          :        :  :|      .  .  |.    ...    |   .      ...|      .      :    | :
Cp    194 GGCTGTCTACATAGGAGGAAGGTATGTATATGGAGCCGTGCTCCAGGGGTAAGATGGATT 135

Db    859 sv tnynsdnvantanvanstnrnvvtnnndnytrndanndnndvykvntnqdaymvvsq 917
          :: | : :: :        :      : ::|     : :|   :     :   :  ::    |   : ::: :
Cp    134 GACTGCAGTTGTAGCAGGAAGGACAATTAAGGCTAGATGGTGAGTTTTTTATATCCATCT 75

Db    918 nnqrnqnrhannnarmanannddvssn rnrhrnhdnnrnrnqvhtqnv 965
              : |  :|       | |  : :::  : ::: ::  : :   :: | :
Cp     74 CTTGAGATAACAGCTGAGAAAACACCTTGCAAGAAGAGGCACAAAGGTC 26


RESULT   10
ID   T42193 standard; DNA; 3460 BP.
AC   T42193.
DT   20 JAN 1997  (first entry)
DF   Fragment of pig oestrogen receptor gene.
KW   Pig; oestrogen receptor; exon; polymorphism; allele; RFLP; litter size;
KW   restriction fragment length polymorphism, Chinese, American, PCR, primer,
KW   amplification; polymerase chain reaction; ds.
OS   Sus scrofa
FH   Key              Location/Qualifiers
FT   misc_feature     303..304
FT   /*tag=  a
FT   /note= "undetermined length gap missing in DNA sequence
FT   between these nucleotides"
FT   exon             534..600
FT   /*tag=  b
FT   /note= "exon 3"
FT   misc_feature     1197..1198
FT   /*tag=  c
FT   /note= "approx. 100 200 bases missing from DNA sequence
FT   between these nucleotides"
FT   primer_bind      1898..1915
FT   /*tag=  d
FT   /note= "binds primer A10RII (T42200)"
FT   primer_bind      2011..2031
FT   /*tag=  e
FT   /note= "binds primer DV028 (T42198)"
FT   primer_bind      2146..2171
FT   /*tag=  f
FT   /note= "binds primer ESR REV (T42195)"
FT   primer_bind      2636
FT   /*tag=  g
FT   primer_bind      2636..2663
FT   /*tag=  h
FT   /note= "binds primer ESRSR (T42197)"
FT   variation        2688..2693
FT   /*tag=  i
FT   /note= "polymorphic PvuII site"
FT   primer_bind      2729..2755
FT   /*tag=  j
FT   /note= "binds primer ESRSF (T42196)"
FT   primer_bind      2922..2940
FT   /*tag=  k
FT   /note= "binds primer B1F (T42201)"
FT   primer_bind      2982..3002
FT   /*tag=  l
FT   /note= "binds primer DV024 (T42199)"
FT   primer_bind      3331..3355
FT   /*tag=  m
FT   /note= "binds primer ESR FOR (T42194)"
FT   misc_feature     3123..3128
FT   /*tag=  n
FT   /note= "constant PvuII site"
PN   US5550024-A.
PD   27-AUG-1996.
PF   19-APR-1991; 687708.
PR   19-APR-1991; US-687708.
PR   19-OCT-1992; US-961819.
PR   26-SEP-1994; US-312312.
PA   (BIOT-) BIOTECHNOLOGY RES & DEV CORP.
PA   (IOWA ) UNIV IOWA STATE RES FOUND INC
PI   Jacobson CD,  Mileham AJ,  Plastow GS,  Rothschild MF;
PI   Tuggle CK,  Vaske DA;
DR   WPI; 96-401592/40.
PT   Screening pigs for high litter size - by identifying a polymorphism
PT   in the pig oestrogen receptor gene in a DNA sample
PS   Disclosure; Column 19-24; 26pp; English.
CC   This is a fragment of the pig oestrogen receptor (ESR) gene including
CC   exon 3.  The sequence covers the region contg. a polymorphic PvuII
CC   restriction enzyme site which is used as site for determining the ESR
CC   allele carried by the animal.  Restriction fragment length polymorphism
CC   (RFLP) analysis of this region revealed two alleles of 4.3 and 3.7 kb.
CC   Both alleles are involved with increased pig litter size dependent on the
CC   breed of pig involved e.g. Chinese or American pigs.  However pigs with
CC   the 3.7 kb allele had between 1.78-2.4 more pigs born in each litter and
CC   1.8-2.4 more pigs born alive than pigs with the 4.3 kb allele. The allele
CC   can also be determined by PCR using the primers (T42194-T42201).
SQ   Sequence   3460 BP;    1052 A;    765 C;    634 G;    1009 T;

  Query Match              3.7%,  Score 71,  DB 24,  Length 3460;
  Best Local Similarity  83.3%;  Pred. No. 5.20e-31;
  Matches    95;  Conservative    0;  Mismatches  18;  Indels    1;  Gaps    1;

Db    490 ttcatcatgcccacttcgtagcacttgcgtagccggcagg-tgacagctcttcctccta 548
          |||| ||| ||||||||||| ||||| || || ||||||||| || ||||||||  : ::
Cp    723 TTCACCATTCCCACTTCGTAACACTTCCGAAGTCGGCAGGCTGGCAGCTCTTGCGCCGG 664

Db    549 ttcttatcaattgtgcactggttggtagctggacacatgtagtcattatatcct 602
          || ||||| ||||| ||||| || ||||||||||| || || ||||||||||||
Cp    663 TTTTTATCGATTGTACACTGATTTGTAGCTGGACAAATATAATCATTATGTCCT 610


RESULT   11
```

```
ID   N80922 standard; cDNA; 2153 BP.
AC   N80922;
DT   29-DEC-1990  (first entry)
DE   Sequence encoding human estrogen-related receptor protein (hERR2)
KW   Hormone receptor; hormone-binding; transcription activation; ss.
OS   Homo sapiens.
FH   Key              Location/Qualifiers
FT   CDS              100..1401
FT   /*tag=  a
PN   WO8803168-A.
PD   05-MAY-1988.
PF   22-OCT-1987; U02782
PR   4-OCT-1986; US-922585
PR   20-OCT-1987; US-108471.
PA   (SALK) Salk Inst for Biol Stud.
PI   Evans RM, Weinberger CA, Hollenberg SM  Giguere V;
DR   WPI; 88-133242/19.
DR   P-PSDB; P80931.
PT   Recombinant DNA encoding hormone recpetors -
PT   comprising glucocorticoid  mineralocorticoid, thyroid hormone
PT   and novel hormone receptors
PS   Claim 16; Fig V-2(B)-1 and -2; 243pp; English.
CC   DNAs encoding hormone receptors and the hormone receptors themselves are
CC   claimed  The DNA can be used to make the hormone receptor proteins and
CC   functional modified forms in quantities not previously possible. The
CC   receptor proteins can be used to screen cpds. for receptor-agonist or
CC   receptor-antagonist activity. They can also be used in diagnostic assays.
CC   Also claimed is pure DNA encoding oestrogen-related receptors hERR1 and
CC   hERR2 and hERR1 and hERR2  The new hERR1 and hERR2 receptors will provide
CC   the basis for development of an assay system that will lead to the
CC   identification of novel hormones.
SQ   Sequence   2153 BP;   445 A;   627 C;   624 G;   457 T;

  Query Match              3.6%;  Score 68;  DB 2;  Length 2153;
  Best Local Similarity  67.2%;  Pred  No  6.54e-29;
  Matches   133;  Conservative    0;  Mismatches  65;  Indels   0;  Gaps   0;

Db     412 gtgtgcggggacattgcttctggctaccactatggagtggcctcctgcgaggcttgcaag 471
           || ||| | ||    || || || || |||||||||||    || || || |  || |||
Qy     527 GTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATGTAAG 586

Db     472 gcgttcttcaagagaaccattcaaggaaacatcgaatacagctgccctgccaccaacgag 531
           || || || || |||| |||||||||| | |  || || |  || || || || ||  ||
Qy     587 GCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCAG 646

Db     532 tgtgagatcaccaaacggaggcgcaagtcctgtcaggcctgccggttcatgaaatgcctc 591
           |||   |||   |||    ||||||||  ||| |||||||||||  |   ||| ||   |
Qy     647 TGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTAC 706

Db     592 aaagtggggatgctgaag 609
            ||||||| ||| |||||
Qy     707 GAAGTGGGAATGGTGAAG 724


RESULT   12
ID   Q76269 standard; DNA, 6244 BP.
AC   Q76269;
DT   17-JUL-1995  (first entry)
DE   PhCMV*-1-controlled progesterone receptor
KW   tTA; transactivator; tetracycline-controllable transactivator,
KW   conditional inactivation; homologous recombination; gene expression;
KW   gene regulation; gene therapy; tetracycline-resistance, tetR,
KW   transgenic animal, PhCMV* 1, promoter, tetO, CMV,
KW   progesterone receptor; ds.
OS   Human cytomegalovirus Towne
PN   WO9429442-A.
PD   22-DEC-1994
PF   14-JUN-1994; U06734.
PR   14-JUN-1993; US-076327
PA   (BADI ) BASF AG.
PI   Bujard H,  Gossen M,  Salfeld JG,  Voss JW,
DR   WPI; 95-036472/05.
PT   Regulatory systems using tetracycline-controllable transactivator
```

```
PT   (tTA)  - useful for conditional inactivation or modulation of
PT   gene expression in a host cell or animal
PS   Disclosure; Page 54-58; 103pp; English.
CC   The minimal promoter PhCMV*-1 is based on the human CMV
CC   intermediate-early promoter (with enhancer region removed), fused
CC   to a tetracycline operator, tetO.  The promoter is used to
CC   express tTA transactivator in host cells.  In clone pUHD BGR3
CC   (given in Q76269), cDNA for rabbit progesterone receptor is under
CC   the control of PhCMV*-1.
SQ   Sequence   6244 BP;    1483 A;    1749 C;    1610 G;    1402 T;

  Query Match              3.3%;  Score 63;  DB 15;  Length 6244;
  Best Local Similarity  66.3%;  Pred. No. 1.92e-25;
  Matches   128;  Conservative    0;  Mismatches  65;  Indels   0;  Gaps   0;

Db    2365 tctgtggggatgaagcatcaggctgtcattatggtgtcctcacctgtgggagctgtaagg 2424
           ||||  | ||| | ||||| || | ||| ||||| |||    | ||||   | |||||||
Qy     528 TCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATGTAAGG 587

Db    2425 tcttctttaaaagggcaatggaagggcagcataactatttatgtgctggaagaaatgact 2484
            ||| ||||||||    ||  |||| ||  || | ||| | ||| | |  | |||| | |
Qy     588 CCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCAGT 647

Db    2485 gcattgttgataaaatccgcaggaaaaactgcccggcgtgtcgccttagaaagtgctgtc 2544
           | |   | ||||||| |||  | || | ||||| ||| || || ||| | ||||| |
Qy     648 GTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTACG 707

Db    2545 aagctggcatggt 2557
           |||  || |||||
Qy     708 AAGTGGGAATGGT 720


RESULT   13
ID   T06872 standard, DNA, 6244 BP.
AC   T06872,
DT   13-MAR-1996  (first entry)
DE   PhCMV*-1 promoter and rabbit progesterone receptor.
KW   Transactivator; tTA; tet operator; tetO; gene expression;
KW   tetracycline-responsive promoter; PhCMV*-1; human cytomegalovirus;
KW   transcription activation, eukaryotic cell, progesterone receptor;
KW   ds; ss.
OS   Chimeric synthetic;
OS   Chimeric Oryctolagus cuniculus.
PN   US5464758-A.
PD   07-NOV-1995.
PF   14-JUN-1993, 076726.
PR   14-JUN-1993; US-076726.
PA   (BUJA/) BUJARD H.
PA   (GOSS/) GOSSEN M.
PI   Bujard H,  Gossen M;
DR   WPI; 95-392612/50.
PT   Polynucleotide encoding trans.activator fusion protein contg. tet
PT   repressor  - used to control expression of gene regulated by minimal
PT   promoter linked to tet operon, and vectors and cells where gene
PT   expression is regulated by tetracycline
PS   Disclosure; Fig 9A-C; 37pp; English.
CC   cDNA coding for the rabbit progesterone receptor was placed under
CC   control of the PhCMV*-1 promoter (see T06869), resulting in
CC   tetracycline-regulated expression vector pUHDpgr-3 contg. the
CC   sequence given in T06872.  The construct was transfected into
CC   eukaryotic cells expressing the tTA transactivator (T06867) to
CC   allow tetracycline-regulated expression of the receptor.
SQ   Sequence   6244 BP;    1483 A;    1750 C;    1609 G;    1402 T;

  Query Match              3.3%;  Score 63;  DB 17;  Length 6244;
  Best Local Similarity  66.3%;  Pred. No. 1.92e-25;
  Matches   128;  Conservative    0;  Mismatches  65;  Indels   0;  Gaps   0;

Db    2365 tctgtggggatgaagcatcaggctgtcattatggtgtcctcacctgtgggagctgtaagg 2424
           ||||  | ||| | ||||| || | ||| ||||| |||    | ||||   | |||||||
Qy     528 TCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATGTAAGG 587

Db    2425 tcttctttaaaagggcaatggaagggcagcataactatttatgtgctggaagaaatgact 2484
```

```
               ||| ||||||||     ||   |||| ||   || | ||| | ||| | |   | |||| | |
Qy      588 CCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCAGT 647

Db     2485 gcattattgataaaatccgcaggaaaact [illegible] 2544
             | |   | |.||||| |||   | || | ||||| ||| || || ||| | ||||| |
Qy      648 GTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTACC 707

Db     2545 aagctggcatggt 2557
            |||  || |||||
Qy      708 AAGTGGGAATGGT 720


RESULT    14
ID   Q10388 standard; cDNA; 1912 BP.
AC   Q10388;
DT   11-APR-1991  (first entry)
DE   Encodes skin-specific murine Retinoic Acid Receptor-gamma.
KW   mouse; retinoic acid receptor; mRAR-gamma; embryogenesis; ss.
OS   Mus musculus.
FH   Key                 Location/Qualifiers
FT   CDS                 316..1892
FT   /*tag=  a
FT   /product= murine RAR gamma
PN   EP-411323-A.
PD   06-FEB-1991.
PF   29-JUN-1990; 112469.
PR   30-JUN-1989; US-374690.
PR   29-MAR-1990; US-502140.
PA   (INRM ) INSERM INST NAT SANTE.
PA   (SQUI ) LES LABS SQUIBB SA.
PI   Chambon P, Kastner P, Krust A, Petkovich M, Zelent A,
PI   Leroy P, Mendelsoh C, Staub A;
DR   WPI; 91-038271/06.
DR   P-PSDB; R10549.
PT   Novel human and mouse retinoic acid receptors - encode proteins
PT   used to assay for agonists and antagonists
PS   Disclosure; Fig 1; 33pp; English.
CC   An 11.5 day-old total mouse embryo lambda gt10 cDNA library was
CC   screened with labelled human RAR-alpha and -beta probes. Positive
CC   clones were isolated, mapped and sequenced. A set of clones was
CC   identified having a lesser degree of homology with either RAR-alpha
CC   or -beta than murine RAR-alpha or mRAR-beta clones, respectively
CC   The cDNA-deduced amino acid sequence of this third set of clones
CC   was obviously related to mRAR-alpha and mRAR-beta. This new member
CC   of the mouse RAR subfamily was designated RAR gamma. The
CC   mRAR gamma cDNA clone was used to clone 7 different murine
CC   RAR-gamma cDNA isoforms, generated by alternative splicing of at
CC   least 7 exons. The isoforms were found to differ in their 5' UTR
CC   sequences and in their N-terminal A region, which is known to be
CC   important for differential trans-activation by other nuclear
CC   receptors.
CC   See also R10547-8, Q10389 and Q10405-8
SQ   Sequence   1912 BP;    427 A;    597 C;    534 G;    354 T;

  Query Match             3.1%;  Score 58;  DB 2;  Length 1912;
  Best Local Similarity   64.6%;  Pred. No. 5.08e-22;
  Matches  128;  Conservative    0;  Mismatches   70;  Indels    0;  Gaps    0;

Db      583 tgctttgtatgcaatgacaagtcttctggctaccactatggggtcagctcctgtgaaggc 642
            |||  ||| ||||  ||  |  | || || || |||||||| ||| | || ||||||||
Qy      521 TGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGA 580

Db      643 tgcaagggcttcttcagacgcagcattcagaaaaacatggtgtatacatgtcaccgtgac 702
            || |||| ||| || | | | ||||||||   | | |  |  ||||  ||||    |
Qy      581 TGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACA 640

Db      703 aaaaactgtatcatcaacaaggtcaccagaaatcgatgccagtactgcaggctacaaaag 762
            ||  | ||||  ||| | ||   |    | ||  | ||||||  |||| | || |  |||
Qy      641 AATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAG 700

Db      763 tgtttcgaagtgggcatg 780
            |||| ||||||||| |||
Qy      701 TGTTACGAAGTGGGAATG 718
```

```
RESULT    15
ID   Q20542 standard; DNA; 2521 BP.
AC   Q20542;
DT   22-APR-1992  (first entry)
DE   Mouse retinoic acid receptor isoform clone lambda-mE7.1.
KW   RAR-gamma-B; ss.
OS   Mus musculus.
FH   Key                 Location/Qualifiers
FT   CDS                 67..1410
FT   /*tag=  a
FT   polyA_signal        2478..2483
FT   /*tag=  b
FT   polyA_site          2497..2521
FT   /*tag=  c
FT   misc_feature        1..216
FT   /*tag=  d
FT   /note= "point of divergence from RAR-gamma-B"
PN   CA2015766-A.
PD   31-OCT-1991.
PF   30-APR-1990; 015766.
PR   30-APR-1990; CA-015766.
PA   (MOUN ) MOUNT SINAI HOSPITAL.
PA   (HSCR-) HSC RES DEV CORP.
PI   Giguere V, Varmuza SL, Rossant J;
DR   WPI; 92-024769/04.
DR   P-PSDB; R20466.
PT   DNA segments to produce iso-forms and chimeric forms of RAR -
PT   have sequence encoding protein which display ligand binding and
PT   transcription activating properties of a gamma iso-form of RAR
PS   Claim 6; Fig 2; 52pp; English.
CC   The sequence is that of a clone lambda-mE7.1 which encodes mouse
CC   retinoic acid receptor (RAR) isoform RAR-gamma-B. it was isolated
CC   from a 12.5 day mouse embryo lambda-gt10 cDNA library. It diverges
CC   from the clone lambda mE13.20 as indicated at nucleotides 1 to 216,
CC   this corresponds to with a splice junction between exon 1 and exon
CC   2 of the human RAR-alpha gene. It can be used for screening of
CC   agonists and antagonists of the gamma receptor effector system and
CC   for identifying ligands capable of binding to a gamma isoform of RAR.
CC   See also Q20541.
SQ   Sequence   2521 BP;    557 A;    752 C;    684 G;    528 T;

  Query Match             3.1%;  Score 58;  DB 3;  Length 2521;
  Best Local Similarity   64.6%;  Pred. No. 5.08e-22;
  Matches  128;  Conservative    0;  Mismatches   70;  Indels    0;  Gaps    0;

Db      301 tgctttgtatgcaatgacaagtcttctggctaccactatggggtcagctcctgtgaaggc 360
            |||  ||| ||||  ||  |  | || || || |||||||| ||| | || ||||||||
Qy      521 TGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGA 580

Db      361 tgcaagggcttcttcagacgcagcattcagaaaaacatggtgtatacatgtcaccgtgac 420
            || |||| ||| || | | | ||||||||   | | |  |  ||||  ||||    |
Qy      581 TGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACA 640

Db      421 aaaaactgtatcatcaacaaggtcaccagaaatcgatgccagtactgcaggctacaaaag 480
            ||  | ||||  ||| | ||   |    | ||  | ||||||  |||| | || |  |||
Qy      641 AATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAG 700

Db      481 tgtttcgaagtgggcatg 498
            |||| ||||||||| |||
Qy      701 TGTTACGAAGTGGGAATG 718


Search completed: Thu Dec 18 20:50:57 1997
Job time : 214 secs.
```

```
**********************************************************************

                      MPSRCH                                  (TM)

**********************************************************************

           Release 2.1D John F. Collins, Biocomputing Research Unit
           Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
                  Distribution rights by IntelliGenetics, Inc.

MPsrch_pp   protein - protein database search, using Smith-Waterman algorithm

Run on:           Thu Dec 18 09:04:45 1997   MasPar time 15.53 Seconds
                                             777.769 Million cell updates/sec
Tabular output not generated.

Title:            >US-08-826-361A-21
Description:      (1-418) from US08826361A.pep
Perfect Score:    3116
Sequence:         1 MNYSIPSNVTNLEGGPGRQT          SMPLANLLMLLSHVRHARSA 418

Scoring table:    PAM 150
                  Gap 11

Searched:         91006 seqs, 28889323 residues

Post-processing:  Minimum Match 0%
                  Listing first 45 summaries

Database:         pir51
                    1:ann1 2:ann2 3:ann3 4:ann4 5:unann1 6:unann2 7:unann3
                    8:unann4 9:unann5 10:unann6 11:unann7 12:unann8
                    13:unann9 14:unann10 15:unenc 16:unrev

Statistics:       Mean 47.491;  Variance 101.275;  scale 0.469

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                SUMMARIES

                  %
Result          Query
  No.   Score   Match Length DB  ID           Description           Pred. No.
------------------------------------------------------------------------------
    1    1512    48.5     586  3  QRXLE        estrogen receptor -    4.09e-263
    2    1484    47.6     589  3  QRCHE        estrogen receptor -    1.39e-257
    3    1471    47.2     701 16  S64737       BOK estrogen recepto   5.11e-255
    4    1463    47.0     595  3  QRHUE        estrogen receptor -    1.94e-253
    5    1455    46.7     599  3  QRMSE        estrogen receptor -    7.34e-252
    6    1448    46.5     600  3  QRRTE        estrogen receptor -    1.77e-250
    7    1327    42.6     574  7  A37197       estrogen receptor -    1.21e-226
    8    1314    42.2     595  7  S58224       oestrogen receptor -   4.37e-224
    9     740    23.7     433 14  S58087       estrogen receptor re   1.76e-112
   10     736    23.6     423 12  B29345       steroid hormone rece   1.03e-111
   11     727    23.3     721 12  A29345       steroid hormone rece   1.33e-109
   12     681    21.9     933  7  A33596       progesterone recept.   3.14e-101
   13     671    21.5     923 14  I52290       progesterone recept.   2.50e-99
   14     666    21.4     933  3  QRHUP        progesterone recepto   2.22e-98
   15     662    21.2     930  7  A25922       progesterone recept.   1.29e-97
   16     614    19.7     742  3  QRHUGB       glucocorticoid recep   1.54e-88
   17     614    19.7     777  3  QRHUGA       glucocorticoid recep   1.54e-88
   18     608    19.5     776  3  S44047       glucocorticoid recep   2.09e-87
   19     597    19.2     786  7  A35466       progesterone recepto   2.47e-85
   20     594    19.1     771  7  A54273       glucocorticoid recep   9.05e-85
   21     586    18.8     783  3  A25691       glucocorticoid recep   2.89e-83
   22     579    18.6     133  7  S26595       estrogen receptor -    5.97e-82
   23     577    18.5     795  3  QRRTG        glucocorticoid recep   1.42e-81
   24     562    18.0     981 14  A41401       mineralocorticoid re   9.21e-79
   25     557    17.9     899 14  A35895       androgen receptor -    7.95e-78
   26     556    17.8     902 14  B40494       androgen receptor -    1.22e-77
   27     550    17.7     910 13  A34721       androgen receptor A    1.62e-76
   28     550    17.7     911 13  B34721       androgen receptor B    1.62e-76
   29     553    17.7     919 13  A39248       androgen receptor -    4.45e-77
   30     544    17.5     758 16  S55586       glucocorticoid recep   2.14e-75
   31     540    17.3     344 13  I51330       androgen receptor -    1.19e-74
   32     523    16.8     984 12  A29513       mineralocorticoid re   1.76e-71
   33     505    16.2     410 14  S26669       retinoic acid recept   3.88e-68
   34     505    16.2     446 14  A24418       H-2 region II bindin   3.88e-68
   35     505    16.2     448 14  D41717       retinoid X receptor    3.88e-68
   36     505    16.2     451 14  A41651       retinoic acid recept   3.88e-68
   37     506    16.2     533 13  S47781       retinoid X receptor    2.53e-68
   38     501    16.1     563  7  JQ0623       nerve growth factor-   2.14e-67
   39     494    15.9     466 14  C41717       retinoid X receptor    4.23e-66
   40     494    15.9     467 14  S26668       retinoic acid recept   4.23e-66
   41     494    15.9     467 14  A47278       retinoid X receptor    4.23e-66
   42     495    15.9     598  7  A37251       probable nuclear hor   2.76e-66
   43     493    15.8     462 13  S09542       retinoic acid recept   6.48e-66
   44     490    15.7     598  7  A61451       steroid hormone rece   2.33e-65
   45     489    15.7     601  3  QRMSN1       probable hormone rec   3.36e-65

                                ALIGNMENTS

RESULT    1
ENTRY              QRXLE      #type complete
TITLE              estrogen receptor - African clawed frog
ORGANISM           #formal_name Xenopus laevis #common_name African clawed frog
DATE               29-Feb-1992 #sequence_revision 14-Jul-1994 #text_change
                     15-Jun-1996
ACCESSIONS         A40907
REFERENCE          A40907
   #authors        Weiler, I.J.; Lew, D.; Shapiro, D.J.
   #journal        Mol. Endocrinol. (1987) 1:355-362
   #title          The Xenopus laevis estrogen receptor: sequence homology with
                     human and avian receptors and identification of multiple
                     estrogen receptor messenger ribonucleic acids.
   #cross-references MUID:90331927
   #accession      A40907
      ##molecule_type mRNA
      ##residues        1-586 ##label WEI
      ##cross-references GB:L20735
COMMENT    The steroid hormones and their receptors are involved in the
             regulation of eukaryotic gene expression and affect cellular
             proliferation and differentiation in target tissues.
COMMENT    In the absence of ligand, steroid hormone receptors are thought to
             be weakly associated with nuclear components; hormone binding
             greatly increases receptor affinity. The hormone-receptor complex
             appears to recognize discrete DNA sequences upstream of
             transcriptional start sites.
CLASSIFICATION     #superfamily estrogen receptor, erbA transforming protein
                     homology
KEYWORDS           DNA binding; nucleus; phosphoprotein; steroid binding;
                     steroid hormone receptor; transcription regulation; zinc
                     finger
FEATURE
   1-174               #domain amino-terminal #label NH2\
   175-266             #domain DNA binding #status predicted #label DNA\
   178-448             #domain erbA transforming protein homology #label ERBA\
   178-201             #region zinc finger C4C4 motif\
   214-237             #region zinc finger C2C2 motif\
   251-266             #region nuclear location signal\
   292-544             #domain steroid binding #status predicted #label STE\
   180,183,197,200     #binding_site zinc (Cys) #status predicted\
   216,222,232,235     #binding_site zinc (Cys) #status predicted\
   231                 #binding_site phosphate (Ser) (covalent) #status
                         predicted
SUMMARY            #length 586  #molecular weight 66080  #checksum 9991
```

```
  Query Match             48.5%,  Score 1512,  DB 3,  Length 586,
  Best Local Similarity   52.1%;  Pred  No  4 09e-263;
  Matches  213;  Conservative   95;  Mismatches  84;  Indels  17;  Gaps  13;

Db      101 spvvflaklpqlspfihhhgqqvpyyleseqqtf-avreaapptfyrsssdnrrqsqrer 159
            || |: :   :|||:: |: |     | |:: :         |            : :|: || :|
Qy       22 SPNVLWPTPGHLSPLVVHR-QLSHLYAEPQKSPWCEARSLEHTLPVNRETLKRKVSG-NR 79

Db      160 mssandkqppsmestketrycavcsdyasgyhygvwscegckaffkrsiqghndymcpat 219
              ::    || :   | :::::|||||||||||||||||||||||||||||||||||:||||
Qy       80  CASPVIGP-G--SKPDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAT 135

Db      220 nqctidknrrksrqacrlrkcyevgmmkggirkdrrggrllkhkrqkeeqeqkndvdpse 279
            ||||||||||||||||||||||||||:| | |::| | ||::  ||.   .|| .
Qy      136 NQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVR--RQRSADEQLH-CAGKA 192

Db      280 irtasiwvnpsvksmklspvlsltaeqlisalmeaeapivysehdstkplseasmmtlit 339
             |:::     | |: : | : ||  :|||: :|:|||:| |    : |: |::|||||  ||
Qy      193 KRSGG--HAPRVRFLILDA-LS--PEQLVLTLLEAEPPHVLISRPSA-PFTEASMMMSLT 246

Db      340 nladrelvhminwakrvpqfvdltlhdqvhllecawleilmvgliwrsvehpgklsfapn 399
            :|||:|||||||:|||::||||:|:| |||:|||  |:|:||:||:|||::||||| |||:
Qy      247 KLADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPD 306

Db      400 llldrnqqrcveqlveifdmlvttatrfrmmrlrqeeficlksiillnsgvytflsstle 459
            |:||||::|:||||:|||||  |  |||  | |   |  : :  :||||:.|  ....: .
Qy      307 LVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQD 366

Db      460 sledtdlihiildkiidtlvhtmaksqlslqqqrrlaqllliilshirh 508
             :       | : |: : |:||  :||||:| |||  ||| ||::|||:||
Qy      367 ADSSRKLAHL-LNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRH 414
```

```
RESULT      2
ENTRY              ORCHE       #type complete
TITLE              estrogen receptor - chicken
ORGANISM           #formal_name Gallus gallus #common_name chicken
DATE               06-Mar-1992 #sequence_revision 14-Jul-1994 #text_change
                     15-Jun-1996
ACCESSIONS         A40914; S07192
REFERENCE          A40914
   #authors        Maxwell, B.L.; McDonnell, D.P.; Conneely, O.M.; Schulz, T.Z.;
                     Greene, G.L.; O'Malley, B.W.
   #journal        Mol. Endocrinol. (1987) 1:25-35
   #title          Structural organization and regulation of the chicken
                     estrogen receptor.
   #cross-references MUID:88318621
   #accession      A40914
      ##status          not compared with conceptual translation
      ##molecule_type mRNA
      ##residues        1-589 ##label MAX
REFERENCE          S07192
   #authors        Krust, A.; Green, S.; Argos, P.; Kumar, V.; Walter, P.;
                     Bornert, J.M.; Chambon, P.
   #journal        EMBO J. (1986) 5:891-897
   #title          The chicken oestrogen receptor sequence: homology with v-erbA
                     and the human oestrogen and glucocorticoid receptors
   #cross-references MUID:86247579
   #accession      S07192
      ##status          preliminary
      ##molecule_type mRNA
      ##residues        1-256,'E',258-589 ##label KRU
      ##cross-references EMBL:X03805
COMMENT     The steroid hormones and their receptors are involved in the
              regulation of eukaryotic gene expression and affect cellular
              proliferation and differentiation in target tissues
COMMENT     In the absence of ligand, steroid hormone receptors are thought to
              be weakly associated with nuclear components. hormone binding
              greatly increases receptor affinity  The hormone-receptor complex
              appears to recognize discrete DNA sequences upstream of
              transcriptional start sites
CLASSIFICATION   #superfamily estrogen receptor; erbA transforming protein
                   homology
KEYWORDS           DNA binding, nucleus, phosphoprotein; steroid binding;
                     steroid hormone receptor, transcription regulation; zinc
                     finger
FEATURE
   1-173               #domain amino-terminal #label NH2\
   174-265             #domain DNA binding #status predicted #label DNA\
   177-450             #domain erbA transforming protein homology #label ERBA\
   177-200             #region zinc finger CCCC motif\
   213-235             #region zinc finger CCCC motif\
   250-265             #region nuclear location signal\
   294-546             #domain steroid binding #status predicted #label STB\
   179,182,196,199     #binding_site zinc (Cys) #status predicted\
   215,221,231,234     #binding_site zinc (Cys) #status predicted\
   230,299             #binding_site phosphate (Ser) (covalent) #status
                         predicted
SUMMARY            #length 589  #molecular-weight 66773  #checksum 793
```

```
  Query Match             47.6%;  Score 1484;  DB 3;  Length 589;
  Best Local Similarity   52.1%,  Pred. No. 1.39e-257;
  Matches  214;  Conservative   95;  Mismatches  84;  Indels  18;  Gaps  11;

Db      100 spvvflqtapqlspfihhhsqqvpyyleneqqsfqmreaappafyrpssdnrrhsirerm 159
            || |:  |: :|||:: |: |     | | : ::        :   ::    :  :  :: :
Qy       22 SPNVLWPTPGHLSPLVVHR-QLSHLYAEPQKSPWCEARSLEHTL-PVNRETLKRKVSGNR 79

Db      160 astneksslsmestketrycavcsdyasgyhygvwscegckaffkrsiqghndymcpatn 219
              ::    |: :    | :::::||||:|||||||||||||||||||||||||||||:||||
Qy       80 CASPVIGP-G--SKPDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAIN 136

Db      220 qctidknrrkscqacrlrkcyevgmmkggirkdrrggrmmkqkrqreeqdsrnaeasst 279
            |||||||||||||||||||||||||:| | |::| | |:::: |  :||     | |  :
Qy      137 QCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSG 196

Db      280 lraptlwtsplvvkhnkknspalsltaeqmvsalleaeppivyseydpnrpfneasmmtl 339
             .||  ..   |:.      :  |||  :||.| .||||||| |       |. || |||||
Qy      197 GHAPRVR-F-ILL-----D--ALS--PEQLVLTLLEAEPPHVLIS-RPSAPFTEASMMMS 244

Db      340 ltnladrelvhminwakrvpgfvdltlhdqvhllecawleilmiglvwrsmehpgkllfa 399
            ||:|||:||||||||:|||::||||:|:| |||:|||  |:|:||:||:|||::||||:||
Qy      245 LTKLADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFA 304

Db      400 pnllldrnqgkcvegmveifdmllatssrfrmmnlqgeefvclksiillnsgvytflsst 459
            |:|:||||::|||||||::||||||||||::||| ::||  |::|:|::|||||::| ::::|
Qy      305 PDLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTAT 364

Db      460 lksleerdyihrvldkitdtlihlmaksglslqqqhrrlaqllliilshirh 510
              :   |   | :|: :||:|: ::||||:| |||  ||| ||::|||:||
Qy      365 QDADSSRKLAH-LLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRH 414
```

```
RESULT      3
ENTRY              S64737      #type complete
TITLE              80K estrogen receptor - human
ORGANISM           #formal_name Homo sapiens #common_name man
DATE               28-Oct-1996 #sequence_revision 28-Oct-1996 #text_change
                     28-Oct-1996
ACCESSIONS         S64737
REFERENCE          S64737
   #authors        Pink, J.J.; Wu, S.Q.; Wolf, D.M.; Bilimoria, M.M.; Jordan,
                     V.C.
   #journal        Nucleic Acids Res. (1996) 24:962-969
   #title          A novel 80 kDa human estrogen receptor containing a
                     duplication of exons 6 and 7.
   #accession      S64737
      ##status          preliminary
      ##residues        1-701 ##label PIN
      ##cross-references EMBL:U47678
SUMMARY            #length 701  #molecular-weight 78433  #checksum 4951

  Query Match             47.2%;  Score 1471;  DB 16;  Length 701;
  Best Local Similarity   57.1%,  Pred. No. 5.11e-255;
  Matches  196;  Conservative   85;  Mismatches  47;  Indels  15;  Gaps  11;
```

```
Db    178 saketrycavcndyasgyhygvwscegckaffkrsiqghndymcpatnqctidknrrksc 237
          | :::::||||:||||||||||||||:|| |||||||||:||||:||||||||||||| ||||
Qy     89 SKFDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSC 148

Db    238 qacrlrkcyevgmmkggirkdrrggrmlkhkrqrddgegrgevgsagdmraanlwpsplm 297
          ||||||||||||||:| | |::| | |::: |||    | :    :: : :|::    :|
Qy    149 QACRLRKCYEVGMVKCGSRRERCGYRLVR--RQRSADE-QLHCAGKA-KRSGG--HAP-R 201

Db    298 ikrskknslalsltadqmvsalldaepp-ilyseydptrpfseasmmglltnladrelvh 356
          : |      |  :|:::|:| :||:|||| :| |    |: ||:|||||  ||:|||:||||
Qy    202 V-RELL-LD-ALSPEQLVLTLLEAEPPHVLISR--PSAPFTEASMMMSLTKLADKELVH 255

Db    357 minwakrvpgfvdltlhdqvhllecawleilmiglvwrsmehpgkllfapnllldrnqgk 416
          ||:|||::||||::| |||:|||  ::|:||:||:|||::|| ||:|||:|:|||::||
Qy    256 MISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGK 315

Db    417 cvegmveifdmllatssrfrmmnlqgeefvclksiillnsgvytflsstlksleekdhih 476
          ||||::|||||||||:||||  ::||  |::|:|::|||||::| ::::|  : :     :
Qy    316 CVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSRKLA 374

Db    477 rvldkitdtlihlmakagltlqqqhqrlaqlllilshirhmrn 519
          ::|: :|||:|: ::|||::: |||  ||| ||::|||:|| |:
Qy    375 HLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHARS 417


RESULT    4
ENTRY              QRHUE       #type complete
TITLE              estrogen receptor - human
ALTERNATE_NAMES    hER
ORGANISM           #formal_name Homo sapiens #common_name man
DATE               28-May-1986 #sequence_revision 28-May-1986 #text_change
                     15-Nov-1996
ACCESSIONS         A94284; A93376; A43021; S27143; S34000, A41925, B41925,
                     A03244; C41925; D41925
REFERENCE          A94284
   #authors        Greene, G.L.; Gilna, P., Waterfield, M., Baker, A.; Hort, Y.;
                     Shine, J.
   #journal        Science (1986) 231:1150-1154
   #title          Sequence and expression of human estrogen receptor
                     complementary DNA.
   #cross-references MUID:86122927
   #accession      A94284
      ##molecule_type mRNA
      ##residues       1-595 ##label GR1
      ##cross-references GB:M12674
REFERENCE          A93376
   #authors        Green, S., Walter, P., Kumar, V., Krust, A., Bornert, J.M.,
                     Argos, P.; Chambon, P.
   #journal        Nature (1986) 320:134-139
   #title          Human oestrogen receptor cDNA: sequence, expression and
                     homology to v-erb-A.
   #cross-references MUID:86146892
   #accession      A93376
      ##molecule_type mRNA
      ##residues       1-595 ##label GR2
      ##cross-references GB:X03635
REFERENCE          A43021
   #authors        Ponglikitmongkol, M.; Green, S.; Chambon, P.
   #journal        EMBO J. (1988) 7:3385-3388
   #title          Genomic organization of the human oestrogen receptor gene
   #accession      A43021
      ##molecule_type DNA
      ##residues       143-161;205-225;244-264,356-374,'G',376,402-422,447-460,
                         'P',462-467,508-528 ##label PON
REFERENCE          S27140
   #authors        Keaveney, M., Klug, J., Gannon, F.
   #journal        DNA Seq. (1992) 2:347-358
   #title          Sequence analysis of the 5' flanking region of the human
                     estrogen receptor gene.
   #accession      S27143
      ##status         preliminary; translation not shown
      ##molecule_type DNA
      ##residues       1-115 ##label KEA
```

```
      ##cross-references EMBL:X62462
REFERENCE          S34000
   #authors        Pfeffer, U., Fecarotta, E., Castagnetta, L.; Vidali, G.
   #journal        Cancer Res. (1993) 53:741-743
   #title          Estrogen receptor variant messenger RNA lacking exon 4 in
                     estrogen-responsive human breast cancer cell lines.
   #accession      S34000
      ##status         preliminary
      ##molecule_type mRNA
      ##residues       216-254,367-399,'G',401-434 ##label PFE
      ##cross-references EMBL:X73067
      ##note           the authors translated the codon GGG for residue 400 as
                         Val
REFERENCE          A41925
   #authors        Dotzlaw, H.; Alkhalaf, M.; Murphy, L.C.
   #journal        Mol. Endocrinol. (1992) 6:773-785
   #title          Characterization of estrogen receptor variant mRNAs from
                     human breast cancers.
   #cross-references MUID:92293154
   #accession      A41925
      ##molecule_type mRNA
      ##residues       1-214,'ELPTLC' ##label DOT
      ##cross-references NCBIN:106580
      ##experimental_source clone 4; breast cancer
      ##note           sequence has been revised after extraction from NCBI
                         backbone
      ##note           the complete sequence of neither the nucleotide nor the
                         protein is shown in this paper
   #accession      B41925
      ##status         significant sequence differences
      ##molecule_type mRNA
      ##cross-references NCBIN:106597; GB:M69296
      ##experimental_source clone 24; breast cancer
COMMENT    The steroid hormones and their receptors are involved in the
             regulation of eukaryotic gene expression and affect cellular
             proliferation and differentiation in target tissues.
COMMENT    In the absence of ligand, steroid hormone receptors are thought to
             be weakly associated with nuclear components, hormone binding
             greatly increases receptor affinity. The hormone-receptor complex
             appears to recognize discrete DNA sequences upstream of
             transcriptional start sites.
GENETICS
   #gene           GDB:ESR
      ##cross-references GDB:119120
   #map_position 6q25.1-6q25.1
   #introns        151/2; 215/1; 254/1; 366/1; 416/2; 477/1; 518/2
CLASSIFICATION     #superfamily estrogen receptor, erbA transforming protein
                     homology
KEYWORDS           DNA binding, nucleus, phosphoprotein, steroid hormone
                     receptor; transcription regulation; zinc finger
FEATURE
   1-120                 #domain amino-terminal #label NH2\
   121-239               #domain DNA binding #status predicted #label DNA\
   183-456               #domain erbA transforming protein homology #label ERBA\
   185-205               #region zinc finger CCCC motif\
   221-245               #region zinc finger CCCC motif\
   300-595               #domain steroid binding #status predicted #label STB\
   236,305               #binding_site phosphate (Ser) (covalent) #status
                           predicted
SUMMARY            #length 595  #molecular weight [illegible]  #checksum 4317

  Query Match              47.0%;  Score 1463;  DB 3;  Length 595;
  Best Local Similarity    57.4%,  Pred. No. 1.94e-253;
  Matches  195,  Conservative   84,  Mismatches  46,  Indels  15,  Gaps  11;

Db    178 saketrycavcndyasgyhygvwscegckaffkrsiqghndymcpatnqctidknrrksc 237
          | :::::|||||:||||'||||||||:||||||||||||||||||||:|:|||||||||||||
Qy     89 SKFDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSC 148

Db    238 qacrlrkcyevgmmkggirkdrrggrmlkhkrqrddgegrgevgsagdmraanlwpsplm 297
          ||||||||||||||:| | |..| | |... |||    | .    .. . .|::    :|
Qy    149 QACRLRKCYEVGMVKCGSRRERCGYRLVR--RQRSADE-QLHCAGKA-KRSGG--HAP-R 201
```

```
Db    298 ikrskknslal-sltadqmvsalldaepp-ilyseydptrpfseasmmglltnladrelvh 356
           : |      |  :|:::|:| :||:|||| :| |    |: ||:|||||  ||:|||:||||
Qy    202 V-RELLLD-ALSPEQLVLTLLEAEPPHVLISR--PSAPFTEASMMMSLTKLADKELVH 255

Db    357 minwakrvpgfvdltlhdqvhllecawleilmiglvwrsmehpgkllfapnllldrnqgk 416
          ||:|||::||||:|:| |||:|||  |:|:||:||:|||::|| ||:|||:|:|||::||
Qy    256 MISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGK 315

Db    417 cvegmveifdmllatssrfrmmnlqgeefvclksiillnsgvytflsstlksleekdhih 476
          ||||::||||||||||:||||  ::||  |::|:|::|||||::| ::::|  :  :    :
Qy    316 CVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSRKLA 374

Db    477 rvldkitdtlihlmakaqltlqqqhqrlaqlllilshirh 516
          ::|: :||:|: ::||:|:: |||  ||| ||::|||:||
Qy    375 HLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRH 414


RESULT     5
ENTRY              ORMSE      #type complete
TITLE              estrogen receptor - mouse
ORGANISM           #formal_name Mus musculus #common_name house mouse
DATE               06-Mar-1992 #sequence_revision 14-Jul-1994 #text_change
                     15-Jun-1996
ACCESSIONS         A40061
REFERENCE          A40061
   #authors        White, R.; Lees, J.A.; Needham, M.; Ham, J.; Parker, M.
   #journal        Mol. Endocrinol. (1987) 1:735-744
   #title          Structural organization and expression of the mouse estrogen
                     receptor.
   #cross-references MUID:91042558
   #accession      A40061
      ##molecule_type mRNA
      ##residues       1-599 ##label WHI
      ##cross-references GB:M38651
COMMENT    The steroid hormones and their receptors are involved in the
             regulation of eukaryotic gene expression and affect cellular
             proliferation and differentiation in target tissues.
COMMENT    In the absence of ligand, steroid hormone receptors are thought to
             be weakly associated with nuclear components; hormone binding
             greatly increases receptor affinity. The hormone-receptor complex
             appears to recognize discrete DNA sequences upstream of
             transcriptional start sites.
CLASSIFICATION   #superfamily estrogen receptor; erbA transforming protein
                   homology
KEYWORDS          DNA binding; nucleus; phosphoprotein; steroid binding;
                    steroid hormone receptor; transcription regulation; zinc
                    finger
FEATURE
   1-183              #domain amino-terminal #label NH2\
   184-275            #domain DNA binding #status predicted #label DNA\
   187-460            #domain erbA transforming protein homology #label ERBA\
   187-210            #region zinc finger CCCC motif\
   223-247            #region zinc finger CCCC motif\
   260-275            #region nuclear location signal\
   304-556            #domain steroid binding #status predicted #label STB\
   189,192,206,209    #binding_site zinc (Cys) #status predicted\
   225,231,241,244    #binding_site zinc (Cys) #status predicted\
   240,309            #binding_site phosphate (Ser) (covalent) #status
                        predicted
SUMMARY            #length 599  #molecular-weight 66954  #checksum 4749

  Query Match             46.7%;  Score 1455;  DB 3;  Length 599;
  Best Local Similarity   51.7%;  Pred. No. 7.34e-252;
  Matches   214;  Conservative   73;  Mismatches   97;  Indels   29;  Gaps   14;

Db    108 spqlmllhpppqlspflqphgqqvpyylenepsay-tvrdtgppafyrsnsdnrrqngr 166
           : || :|    | :||| |  | |      | | : |:       |           :  :|: :|
Qy     20 TTSPNVLWPTPGHLSP LVVHRQLSHLYAEPQKSPWCEARSLEHTLPVNRETLKRKVSG- 77

Db    167 erlsssnekgnmimesaketrycavcndyasgyhygvwscegckaffkrsiqghndymcp 226
           :|  :|     |     | :::::||||:||||||||||||||||||||||||||||||:||
Qy     78 NR-CASPVTGP---GSKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICP 133
```

```
Db    227 atnqctidknrrksCqacrlrkcyevgmmkggirkdrrggrmlkhkrqrddlegrnemga 286
          ||||||||||||||||||||||||||||||||:| | |::| | |:::  |||   : :    !
Qy    134 ATNQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVR--RQRSA-DEQLH-CA 189

Db    297 sqdmrranlwpsplvikhtkknspalsltadqmvsalldaeppmiyseydpsrpfseasm 346
           .   .|..     .| | .      .  |||   ..|.| :||:|||| :      || ||:||||
Qy    190 GKAKRSGG--HAPRV-RELLLD--ALS--PEQLVLTLLEAEPPHVLIS-RPSAPFTEASM 241

Db    347 mglltnladrelvhminwakrvpgfgdlnlhdqvhllecawleilmiglvwrsmehpgkl 406
          |  ||:|||:||||||:|||::|||  :|:| |||:|||  |:|:||:||:|||::|||||
Qy    242 MMSLTKLADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKL 301

Db    407 lfapnllldrnqgkcvegmveifdmllatssrfrmmnlqgeefvclksiillnsgvytfl 466
          :|||:|:|||::||||||::||||||||||:||||  ::||  |::|:|::|||||::| ::
Qy    302 IFAPDLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLV 361

Db    467 sstlksleekdhihrvldkitdtlihlmakaqltlqqqhrrlaqlllilshirh 520
          ::|  : :    : ::|: :||:|: ::||:|:: |||  ||| ||::|||:||
Qy    362 TATQDA-DSSRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRH 414


RESULT     6
ENTRY              ORRTE      #type complete
TITLE              estrogen receptor - rat
ORGANISM           #formal_name Rattus norvegicus #common_name Norway rat
DATE               30-Sep-1991 #sequence_revision 14-Jul-1994 #text_change
                     15-Jun-1996
ACCESSIONS         S07379; S16731
REFERENCE          S07379
   #authors        Koike, S.; Sakai, M.; Muramatsu, M.
   #journal        Nucleic Acids Res. (1987) 15:2499-2513
   #title          Molecular cloning and characterization of rat estrogen
                     receptor cDNA.
   #cross-references MUID:87174780
   #accession      S07379
      ##molecule_type mRNA
      ##residues       1-600 ##label KOI
      ##cross-references EMBL:Y00102
REFERENCE          S16731
   #authors        Maggi, A.M.A.
   #submission     submitted to the EMBL Data Library, June 1991
   #accession      S16731
      ##molecule_type mRNA
      ##residues       1-487,'T',489-600 ##label MAG
      ##cross-references EMBL:X61098
COMMENT    The steroid hormones and their receptors are involved in the
             regulation of eukaryotic gene expression and affect cellular
             proliferation and differentiation in target tissues.
COMMENT    In the absence of ligand, steroid hormone receptors are thought to
             be weakly associated with nuclear components; hormone binding
             greatly increases receptor affinity. The hormone-receptor complex
             appears to recognize discrete DNA sequences upstream of
             transcriptional start sites.
CLASSIFICATION   #superfamily estrogen receptor; erbA transforming protein
                   homology
KEYWORDS          DNA binding; nucleus; phosphoprotein; steroid binding;
                    steroid hormone receptor; transcription regulation; zinc
                    finger
FEATURE
   1-184              #domain amino-terminal #label NH2\
   185-276            #domain DNA binding #status predicted #label DNA\
   188-461            #domain erbA transforming protein homology #label ERBA\
   188-211            #region zinc finger CCCC motif\
   224-248            #region zinc finger CCCC motif\
   261-276            #region nuclear location signal\
   305-557            #domain steroid binding #status predicted #label STB\
   190,193,207,210    #binding_site zinc (Cys) #status predicted\
   226,232,242,245    #binding_site zinc (Cys) #status predicted\
   241,310            #binding_site phosphate (Ser) (covalent) #status
                        predicted
SUMMARY            #length 600  #molecular-weight 67029  #checksum 4599

  Query Match             46.5%;  Score 1448;  DB 3;  Length 600;
  Best Local Similarity  51.0%;  Pred. No. 1.77e-250;
  Matches  211;  Conservative  97;  Mismatches  96;  Indels  20;  Gaps  14;

Db     109 spsplmllhppphvsrflhphghqvpyyleteqpsay-avrdtqppafyrsnsdnrrqggr 167
            : || :|    | |:|| |  | :     | | : |:    |          :: :|: :|
Qy      20 TTSPNVLWPTPGHLSP-LVVHRQLSHLYAEPQKSPWCEARSLEHTLPVNRETLKRKVSG- 77

Db     168 erlssssekqnmimesaketrycavcndyasgyhygvwscegckaffkrsiqghndymcp 227
            :|  :|:  |        | ::::||||:||||||||||||||||||||||||||||||:||
Qy      78 NR-CASPVTGP---GSKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICP 133

Db     228 atnqctidknrrkscqacrlrkcyevgmmkggirkdrrggrmlkhkrqrddlegrnemgt 287
           |||||||||||||||||||||||||||||||||:| | |::| | |:... |  |. . .  |
Qy     134 ATNQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQRSADE-QLHCA-GK 191

Db     288 sgdmraanlwpsplvikhtkkrspalsltadqmvsalldaepplyseydpsrpfseasm 347
            :  :|::     :| | :     :  |||  ::|:| :||:|||| :      || ||:||||
Qy     192 A--KRSGG--HAPRV-RELLLD--ALS--PEQLVLTLLEAEPPHVLIS-RPSAPFTEASM 241

Db     348 mglltnladrelvhmirwakrvpgfsdlnlhdqvhllecawleilmiglvwrsmehpgkl 407
            |  ||:|||:||||:|||:||::|||  :|:| ||:|||  |:|:||:|||:|||::||||
Qy     242 MMSLTKLADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKL 301

Db     408 lfapnllldrnqgkcvegmveifdmllatssrfrmmnlqgeefvclksiillnsgvytfl 467
           :|||  |  |   |||| |  ||||||||:|||   ||   |  :| | .:||||  :|
Qy     302 IFAPDLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLV 361

Db     468 sstlksleekdhihrvldkindtlihlmakagltlqqqhrrlaqllllilshirh 521
           ::|  :  :    : :::: : |:|: ::||:|:: |||  ||| ||::||||:||
Qy     362 TATQDA-DSSRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRH 414


RESULT     7
ENTRY              A37197      #type complete
TITLE              estrogen receptor - rainbow trout
ORGANISM           #formal_name Oncorhynchus mykiss #common_name rainbow trout
DATE               31-Jan-1992 #sequence_revision 31-Jan-1992 #text_change
                     03-Jun-1996
ACCESSIONS         A37197; A40070
REFERENCE          A37197
   #authors        Pakdel, F.; Le Gac, F.; Le Goff, P.; Valotaire, Y.
   #journal        Mol. Cell. Endocrinol. (1990) 71:195-204
   #title          Full-length sequence and in vitro expression of rainbow trout
                     estrogen receptor cDNA.
   #cross-references MUID:91006824
   #accession      A37197
      ##status          preliminary
      ##molecule_type mRNA
      ##residues        1-574 ##label PAK
      ##cross-references GB:M31559
REFERENCE          A40070
   #authors        Pakdel, F.; Le Guellec, C.; Vaillant, C.; Le Roux, M.G.;
                     Valotaire, Y.
   #journal        Mol. Endocrinol. (1989) 3:44-51
   #title          Identification and estrogen induction of two estrogen
                     receptors (ER) messenger ribonucleic acids in the rainbow
                     trout liver: sequence homology with other ERs.
   #cross-references MUID:89127284
   #accession      A40070
      ##status          preliminary
      ##molecule_type mRNA
      ##residues        150-574 ##label PA2
      ##cross-references GB:M31559
CLASSIFICATION     #superfamily estrogen receptor; erbA transforming protein
                     homology
KEYWORDS           DNA binding; nucleus; steroid hormone receptor; zinc finger
FEATURE
   145-421               #domain erbA transforming protein homology #label ERBA\
   147-167               #region zinc finger\
   183-207               #region zinc finger
SUMMARY            #length 574  #molecular-weight 62971  #checksum 1057

  Query Match             42.6%;  Score 1327;  DB 7;  Length 574;
```

```
  Best Local Similarity  57.5%;  Pred. No. 1.21e-226;
  Matches  196;  Conservative  64;  Mismatches  61;  Indels  20;  Gaps  15;

Db     143 etrycavcsdfasgyhygvwscegckaffkrsiqghndymcpatnqctmdrnrrksqqac 202
           ::::||||||:|||||||||||||||||||||||||||||||:||||||||:::|||||||:
Qy      92 DAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQAC 151

Db     203 rlrkcyevgmvkggirkdrgg-rvlrkdkryvgpaqdrekpytvtwstqqrpqdqdrnsss 261
           |||||||||||| | |::| | |::|:: |   :| | |      : : : :: ||:  ::
Qy     152 RLRKCYEVGMVKCGSRRERCGYRLVRRQ-R---SA-D-EQLHCA-GKAKRS--GGH--AP 200

Db     262 slnggggwrgpritmppeqvlfllqgtp-alesrqkvarpytevtmmtlltsmadkelv 320
               :        : :: | : || ||::: |  | ||   | |:|| :||  || :||||||
Qy     201 PVREL--LLDA-LS-PEQLVLTLLEAEPPHVLISRPS-A-PFTEASMMMSLTKLADKELV 254

Db     321 hmlawakkvpgfqelslhdqvqllesswlevlmiglwwisihcpgklifaqdlildrseq 380
           |||:||||:||| |||| |||:|||| |:||||:||:|||||  |||||| ||:||| ||
Qy     255 HMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEG 314

Db     381 dqvegmaeifdmllatvsrfrmlklkpeefvclkaillnpgafsfcsnsveslhnssav 440
            ||||: ||||||||| |||  |||   |::|:||:||||:: ::: : : :: :|  :
Qy     315 KCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVIATQDA-DSSRKL 373

Db     441 esmldnitdalihhishsqasvqqqprrqaqllllshirh 481
             .|. :|||.  |. || | |||: | | ||:||||:||
Qy     374 AHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRH 414


RESULT     8
ENTRY              S58224      #type fragment
TITLE              oestrogen receptor - Atlantic salmon (fragment)
ORGANISM           #formal_name Salmo salar #common_name Atlantic salmon
DATE               13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change
                     26-Apr-1996
ACCESSIONS         S58224
REFERENCE          S58224
   #authors        Rogers, S.A.; Llewellyn, L.; Ramsurn, V.P.; Sweeney, G.E.;
                     Wigham, T.
   #submission     submitted to the EMBL Data Library, July 1995
   #accession      S58224
      ##status          preliminary
      ##molecule_type mRNA
      ##residues        1-535 ##label ROG
      ##cross-references EMBL:X89959
CLASSIFICATION     #superfamily erbA transforming protein homology
FEATURE
   103-382               #domain erbA transforming protein homology #label ERBA
SUMMARY            #length 535  #checksum 9963

  Query Match             42.2%;  Score 1314;  DB 7;  Length 535;
  Best Local Similarity  56.9%;  Pred. No. 4.37e-224;
  Matches  195;  Conservative  61;  Mismatches  66;  Indels  21;  Gaps  14;

Db     101 etrycavcsdfasgyhygfwscegckaffkrsiqghndymcpatnqctmdrnrrksqqac 160
           ::::||||||:|||||||| ||||||||||||||||||||||:|||||||:|:|||||||||
Qy      92 DAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQAC 151

Db     161 rlrkcyevgmvkggirkdrgg-rvlrkdkryvgpaqdrekpygdlchrtappqdqqrnss 219
           |||||||||||| | |..| | |:.|.. |   :| : .    .     . |  : ||: :
Qy     152 RLRKCYEVGMVKCGSRRERCGYRLVRRQ-R---SADE-QLHCAG---K-AK-RSGGHAPR 201

Db     220 ssslsggggwrgpritmppeqvlfllqgaeppalesrqkvarpytevtmmtlltsmadke 279
              |        : :: | : || ||::  |  | ||   | |:|| :||  || :||||
Qy     202 VREL----L-DA-LS-PEQLVLTLLEAEPPHVLISRPS-A-PFTEASMMMSLTKLADKE 252

Db     280 lvhmlawakkvpgfqelslhdqvqllesswlevlmiglwwisihcpgklifaqdlildrs 339
           |||||:||||:|||| |||| |||:|||| |:||||:||:||||  |||||| ||:|||
Qy     253 LVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRD 312

Db     340 egdcvegmaeifdmllatvsrfrmlklkpeefvclkaillnsgafsfcsnsveslhnss 399
           || ||||. |||||||| |||| |||   |..|.||.|||||. ::: : : :: :|
Qy     313 EGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVIATQDA-DSSR 371
```

```
Db     400 avesmldnitdalihhishsgasvqqqprrqvqlllllshirh 442
              :  :|: :||||:  |: || | |||: |   ||:|||||:||
Qy     372 KLAELLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRH 414


RESULT     9
ENTRY              S58087      #type complete
TITLE              estrogen receptor related protein - mouse
ORGANISM           #formal_name Mus musculus #common_name house mouse
DATE               13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change
                     26-Apr-1996
ACCESSIONS         S58087
REFERENCE          S58087
   #authors        Pettersson, K.; Svensson, K.; Mattsson, R.; Carlsson, B.;
                     Ohlsson, R.; Berkenstam, A.
   #submission     submitted to the EMBL Data Library, July 1995
   #description    Expression of a novel member of estrogen response
                     element-binding nuclear receptors is restricted to the
                     early stages of chorion formation during mouse
                     embryogenesis.
   #accession      S58087
      ##status          preliminary
      ##molecule_type mRNA
      ##residues        1-433 ##label PET
      ##cross-references EMBL:X89594
CLASSIFICATION   #superfamily erbA transforming protein homology
FEATURE
   101-352              #domain erbA transforming protein homology #label ERBA
SUMMARY            #length 433  #molecular weight 48400  #checksum 1075

  Query Match                23.7%;  Score 740;  DB 14;  Length 433;
  Best Local Similarity  42.5%;  Pred. No. 1.76e-112;
  Matches   113;  Conservative   57,  Mismatches  85   Indels  11;  Gaps  10;

Db     101 rlclvcgdiasgyhygvasceackaffkrtiqgnieyncpatneceitkrrrkscqacrf 160
           ::| ||:| |||||||| |||:|||||||:|||: :| |||||:| | | |||||||||:
Qy      94 HFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRL 153

Db     161 mkclkvgmlkegvrldrv---r-qgrqky-krrldsenspylnlpisppakkpltkivsn 215
            ||  |||:| | | :|   |    ||:     :|        :  :    :|  :   |  :| :
Qy     154 RKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPE 213

Db     216 -l-lgv-eqdklyam-ppndipegdikalttlceladrelvfllnwakhipgfpsltlgd 271
            | |  :  | :      : :     |  :    : :|  |||:||| :|:||| ||||  |:| |
Qy     214 QLVLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 273

Db     272 qmsllqsawmeililgivyrslpyddklayaedyimdeehsrlv-glldlyrailqlvrr 330
           |: ||:| |||:|::|:: ||:      || :| | ::| : :: | |:|::: :|     |
Qy     274 QVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSR 333

Db     331 ykklkvekeefmilkalalansdsmy 356
           :: ||::  |:: :||: | || |||
Qy     334 FRELKLQHKEYLCVKAMILLNS-SMY 358


RESULT     10
ENTRY              B29345      #type complete
TITLE              steroid hormone receptor ERR2 precursor - human
ORGANISM           #formal_name Homo sapiens #common_name man
DATE               15-Dec-1988 #sequence_revision 15-Dec-1988 #text_change
                     03-Jun-1996
ACCESSIONS         B29345
REFERENCE          A29345
   #authors        Giguere, V.; Yang, N.; Segui, P.; Evans, R.M.
   #journal        Nature (1988) 331:91-94
   #title          Identification of a new class of steroid hormone receptors.
   #cross-references MUID:88122546
   #accession      B29345
      ##molecule_type mRNA
      ##residues        1-433 ##label GIG
      ##cross-references GB:X51417; GB:Y00290
CLASSIFICATION   #superfamily erbA transforming protein homology
KEYWORDS           DNA binding; zinc finger
FEATURE
   101-352              #domain erbA transforming protein homology #label ERBA\
   103-123              #region zinc finger\
   139-163              #region zinc finger
SUMMARY            #length 433  #molecular-weight 48287  #checksum 1133

  Query Match                23.6%;  Score 736;  DB 13;  Length 433;
  Best Local Similarity  43.6%;  Pred. No. 1.03e-111;
  Matches   116;  Conservative   55;  Mismatches  84;  Indels  11;  Gaps   9;

Db     101 rlclvcgdiasgyhygvasceackaffkrtiqgnieyscpatneceitkrrrkscqacrf 160
           ::| ||:| |||||||| |||:|||||||:|||: :| |||||:| | | |||||||||:
Qy      94 HFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRL 153

Db     161 mkclkvgmlkegvrldrv---r-qgrqky-krrldsenspylslqisppakkpltkivs- 214
            ||  |||:| | | :|   |    ||:     :|       :  |    :|  :  |    :|
Qy     154 RKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPE 213

Db     215 yl-l-vaepdklyam-ppddvpegdikalttlcdladrelvfllswakhipgfsnltlgd 271
            | |  . |..      . .     |   .    . .|  |||.||| .||||| |||| :|:| |
Qy     214 QLVLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 273

Db     272 qmsllqsawmeililgivyrslpyddklayaedyimdeehsrlv-gllelyrailqlvrr 330
           |: ||:| |||:|::|:: ||:      || :| | ::| : :: | |:||::  :|     |
Qy     274 QVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSR 333

Db     331 ykklkvekeefvmlkalalansdsmy 356
           :: ||::  |:: :||: | || |||
Qy     334 FRELKLQHKEYLCVKAMILLNS-SMY 358


RESULT     11
ENTRY              A29345      #type complete
TITLE              steroid hormone receptor ERR1 precursor - human
ALTERNATE_NAMES  estrogen related receptor
ORGANISM           #formal_name Homo sapiens #common_name man
DATE               15-Dec-1988 #sequence_revision 15-Dec-1988 #text_change
                     03-Jun-1996
ACCESSIONS         A29345; A49074
REFERENCE          A29345
   #authors        Giguere, V., Yang, N., Segui, P., Evans, R.M.
   #journal        Nature (1988) 331:91-94
   #title          Identification of a new class of steroid hormone receptors.
   #cross-references MUID:88122546
   #accession      A29345
      ##molecule_type mRNA
      ##residues        1-521 ##label GIG
      ##cross-references EMBL:X51416; EMBL:Y00290
REFERENCE          A49074
   #authors        Wiley, S.R.; Kraus, R.J.; Zuo, F.; Murray, E.E.; Loritz, K.;
                     Mertz, J.E.
   #journal        Genes Dev. (1993) 7:2206-2219
   #title          SV40 early-to-late switch involves titration of cellular
                     transcriptional repressors.
   #accession      A49074
      ##status          preliminary
      ##molecule_type protein
      ##residues        166-169,'X',171-173 ##label WIL
CLASSIFICATION   #superfamily erbA transforming protein homology
KEYWORDS           DNA binding; nucleus; zinc finger
FEATURE
   174-434              #domain erbA transforming protein homology #label ERBA\
   176-196              #region zinc finger\
   212-236              #region zinc finger
SUMMARY            #length 521  #molecular-weight 55742  #checksum 5840

  Query Match                23.3%;  Score 725;  DB 13;  Length 521;
  Best Local Similarity  40.1%;  Pred. No. 1.30e-109;
  Matches   130;  Conservative   67;  Mismatches 117;  Indels  10;  Gaps   7;

Db     174 rlclvcgdvasgyhygvasceackaffkrtiqgsieyscpasneceitkrrrkacqacrf 233
           ::| ||:| |||||||| |||:|||||||:|||  :| ||:|:| | | |||:|||||:
Qy      94 HFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRL 153
```

```
Db    234 tkclrvgmlkegvrldrvrggrqkykrrpevdplpfpgp-fpagplavaggprktaapv 291
           ||   |||:| | |  :|   | |       :     |  :    |         :
Qy    154 RKCYEVGMVKCGSRRERC-GYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSP 212

Db    292 nalvshllvgpeklyampdpagpdghlpavatlcdlfdreivvtiswaksipgfsslsl 351
           : ||   ||  ||   :  :: |::|     : : :|  | | | |  |||| ||||  |||
Qy    213 EQLVLTLLEAEPPHVL-ISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSL 271

Db    352 sdqmsvlqsvwmevlvlgvagrslplqdelafaedlvld-eegaraaglgelgaallqlv 410
           .|: :|:| |||||::|:  ||:      | || ||||| :||     |: |:   ||
Qy    272 FDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATT 331

Db    411 rrlqalrlereeyvllkalalansdsvhiedaeprlwssceklkh--eal--lkyeagrs 465
           |:: |:|.. ||. .||. | ||    :    :    |  || |  :|:    |    :::
Qy    332 SRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKIAHLLNAVTDALVWVIAKS 391

Db    466 gpgggaerrragrllltlpllrqt 489
           | ::  :  | : ||: |: :|::
Qy    392 GISSQQQSMRLANLLMLLSHVRHA 415


RESULT   12
ENTRY              A39596      #type complete
TITLE              progesterone receptor B form - mouse
ORGANISM           #formal_name Mus musculus #common_name house mouse
DATE               20-Mar-1992 #sequence_revision 20-Mar-1992 #text_change
                     03-Jun-1996
ACCESSIONS         A39596
REFERENCE          A39596
   #authors        Schott, D.R.; Shyamala, G.; Schneider, W.; Parry, G.
   #journal        Biochemistry (1991) 30:7014-7020
   #title          Molecular cloning, sequence analyses, and expression of
                     complementary DNA encoding murine progesterone receptor.
   #cross-references MUID:91299759
   #accession    A39596
      ##status          preliminary
      ##molecule_type mRNA
      ##residues        1-923 ##label SCH
      ##cross-references GB:J05333
CLASSIFICATION     #superfamily progesterone receptor; erbA transforming protein
                     homology
KEYWORDS           DNA binding; nucleus; steroid hormone receptor; zinc finger
FEATURE
   555-819              #domain erbA transforming protein homology #label ERBA\
   557-577              #region zinc finger\
   593-617              #region zinc finger
SUMMARY            #length 923  #molecular-weight 99073  #checksum 6627

  Query Match              21.9%;  Score 681;  DB 7;  Length 923;
  Best Local Similarity  30.8%;  Pred. No. 2.14e-101;
  Matches  123;  Conservative  112;  Mismatches 138;  Indels  27;  Gaps  27;

Db    491 aapgtaaapaiyqplg-lnglpqlg-y-qaavlkdslpqvyppylnylrpdseasqspqy 547
           ::||  ::  : : |: | : ||  :    :|  |  :|| |    |:  .  ..   ||
Qy     28 PIPSHLSPLVVHRQLSHLYAFPQKSPWCEARSLEHTIP-VNRETLK-RKVSGNRCASPVT 85

Db    548 gfdslpq-kiclicgdeasgchygvltcgsckvffkramegqhnylcagrndcivdkirr 606
           |  |   :  :| :|:| ||| |||| :| :|| ||||:::|:::|:|:: |:| :|| ||
Qy     86 GPGSKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRR 145

Db    607 kncpacrlrkccqagmvl-ggrkfk-kfnkvrvmrtldgvalpqsvglpnesqalsqrit 664
           |:| ||||||| : |||  |.|.  .   .   ||  |   |   |    | .. |  .  |:
Qy    146 KSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQRSADE-QL-HCAGKAKRSGGHAPRVR 203

Db    665 fspnqeiqlvp-plinllmsiepdviyaghdntkpdtsssllts lnqlgerqllsvvkws 723
                 :  | |  |: |: ||  :  :       | |  |   || |   ::|.    |
Qy    204 -ELLLDA-LSPEQLVLTLLEAEPPHVLISRPSA-PFTEASMMMSLTKLADKELVHMISWA 260

Db    724 kslpgfrnlhiddqitliqyswmslmvfglgwrsykhvsgqmlyfapdliln-eq-r-mk 780
           | :|||  :: : |:: |:: |: ::: || ||| | :| ::  ||||:|:  :: :  :
Qy    261 KKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVE 318
```

```
Db    781 e-lsfyslcltmwqipqefvklqvtheeflcmkvllllnt-i-pleglrsqsqfeemrss 837
           | :: : | :       |  | : | |:||:| ::|||: : ||       :::    : :
Qy    319 GILEIFDM-L-I-ATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAH 375

Db    838 yirelikaig-lrqkgvvptsqrfyqltklldslhdlvkq 876
           :  .  |.  .  |. .:. |.  .|.:||  |    |::
Qy    376 LLNAVTDALVWVIAKSGISSQQQSMRLANLL-MLLSHVRH 414


RESULT   13
ENTRY              I53280      #type complete
TITLE              progesterone receptor - rat
ORGANISM           #formal_name Rattus norvegicus #common_name Norway rat
DATE               02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change
                     02-Aug-1996
ACCESSIONS         I53280
REFERENCE          I53280
   #authors        Park-Sarge, O.K.; Mayo, K.E.
   #journal        Endocrinology (1994) 134:709-718
   #title          Regulation of the progesterone receptor gene by gonadotropins
                     and cyclic adenosine 3',5'-monophosphate in rat granulosa
                     cells.
   #cross-references MUID:94130817
   #accession    I53280
      ##status          preliminary; translated from GB/EMBL/DDBJ
      ##molecule_type mRNA
      ##residues        1-923 ##label RES
      ##cross-references GB:L16922, NID:g463282, CDS_PID:g463283
SUMMARY            #length 923  #molecular-weight 99407  #checksum 9843

  Query Match              21.5%;  Score 671;  DB 14;  Length 923;
  Best Local Similarity  32.6%;  Pred. No. 2.50e-99;
  Matches  102;  Conservative  94;  Mismatches 108;  Indels  21;  Gaps  21;

Db    556 iclicgdeasgchygvltcgsckvffkramegqhnylcagrndcivdkirrkncpacrlr 615
           :| :|:| ||| |||| :| :|| ||||:::|:::|:|:: |:| :|| |||:| |||||
Qy     95 FCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLR 154

Db    616 kccqagmvl-ggrkfk-kfnkvrvmraldgvalpqsvafpnesqtlqqritfspnqeiql 673
           || : |||  |:|: :  :  ||  |: !   | :  : :: |   : |:       :  |
Qy    155 KCYEVGMVKCGSRRERCGYRLVRRQRSADE-QL-HCAGKAKRSGGHAPRVR-ELLLDA-L 210

Db    674 vp-plinllmsiepdvvyaghdntkpdtsssllt slnqlgerqllsvvkwskslpgfrnl 732
           |  |:  |:  ||  |  :: :: | | :|:: ||  |::::|: :: |:| :|||| :|
Qy    211 SPEQLVLTLLEAEPPHVLISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVEL 269

Db    733 hiddqitliqyswmslmvfglgwrsykhvsgqmlyfapdliln-eq-r-mke-lsfyslc 788
           : ||. |..  || .  || ||| |  |     ||||| |   .  .  .   | . . .
Qy    270 SLFDQVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDM- 326

Db    789 ltmwqipqefvklqvtheeflcmkvllllnt-i-pleglrsqsqfeemrssyirelikai 846
           | .      |  | . | |.||.| ..|||. : ||      :::   : :  :  :  |:
Qy    327 L-L-ATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDAL 384

Db    847 g-lrqkgvvpssqrfyqltklldslhdlvkq 876
           .  |. ..| |:  :|::||  |    |::
Qy    385 VWVIAKSGISSQQQSMRLANLL-MLLSHVRH 414


RESULT   14
ENTRY              QRHUP       #type complete
TITLE              progesterone receptor form B - human
ALTERNATE_NAMES    hPR
CONTAINS           progesterone receptor form A
ORGANISM           #formal_name Homo sapiens #common_name man
DATE               30-Jun-1987 #sequence_revision 18-Nov-1994 #text_change
                     15-Nov-1996
ACCESSIONS         S09971; S12464; A03245
REFERENCE          S09971
   #authors        Kastner, P.; Krust, A.; Turcotte, B.; Stropp, U.; Tora, L.;
                     Gronemeyer, H.; Chambon, P.
   #journal        EMBO J. (1990) 9:1603-1614
   #title          Two distinct estrogen-regulated promoters generate
```

```
                 transcripts encoding the two functionally different human
                 progesterone receptor forms A and B
   #cross-references MUID:90228361
   #accession    S09971
      ##molecule_type mRNA
      ##residues       1-933 ##label KAS
      ##cross-references EMBL:X51730
REFERENCE        S12464
   #authors      Kastner, P.
   #submission   submitted to the EMBL Data Library, February 1990
   #accession    S12464
      ##molecule_type mRNA
      ##residues       1-343,'T',345-933 ##label KA2
      ##cross-references EMBL:X51730
REFERENCE        A03245
   #authors      Misrahi, M.; Atger, M.; D'Auriol, L.; Loosfelt, H.; Meriel,
                   C.; Fridlansky, F.; Guiochon-Mantel, A.; Galibert, F.;
                   Milgrom, E.
   #journal      Biochem. Biophys. Res. Commun. (1987) 143:740-748
   #title        Complete amino acid sequence of the human progesterone
                   receptor deduced from cloned cDNA.
   #cross-references MUID:87184565
   #accession    A03245
      ##molecule_type mRNA
      ##residues       1-225,'G',227-255,'V',257-659,'V',661-933 ##label MIS
      ##cross-references GB:M15716
GENETICS
   #gene         GDB:PGR
      ##cross references GDB:119493
   #map_position 11q22.1-11q22.3
CLASSIFICATION   #superfamily progesterone receptor, erbA transforming protein
                   homology
KEYWORDS         alternative splicing; DNA binding; glycoprotein; nucleus;
                   phosphoprotein; steroid hormone receptor; transcription
                   regulation; zinc finger
FEATURE
   1-933               #product progesterone receptor form B #status predicted
                         #label MA1\
   165-933             #product progesterone receptor form A #status predicted
                         #label MA2\
   565-829             #domain erbA transforming protein homology #label ERBA\
   567-587             #region zinc finger CCCC motif\
   603-627             #region zinc finger CCCC motif\
   681-933             #domain steroid binding #status predicted #label STE\
   41                  #binding_site phosphate (Thr) (covalent) #status
                         predicted\
   227,232,552,793     #binding_site phosphate (Ser) (covalent) #status
                         predicted\
   329,374,601         #binding_site phosphate (Tyr) (covalent) #status
                         predicted\
   664                 #binding_site carbohydrate (Asn) (covalent) #status
                         predicted
SUMMARY          #length 933  #molecular-weight 99028  #checksum 5865

  Query Match             21.4%;  Score 666;  DB 3,  Length 933,
  Best Local Similarity   33.4%;  Pred. No. 2.22e-98;
  Matches  110;  Conservative   89;  Mismatches 113;  Indels  17;  Gaps  17;

Db     566 iclicgdeasgchygvltcgsckvffkrameggqhnylcagrndcivdkirrkncpacrlr 625
           :| :|:| ||| |||| :| :|| ||||:::|:::|:|:: |:| :|| |||:| |||||
Qy      95 FCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLR 154

Db     626 kcqqagmvl-ggrkfk-kfnkvrvvraldavalpqplgvpnesqalsqrftfspgqdiql 683
           || : |||  |:|: :   :  ||  |: |    | :  | :: | : : |          | |
Qy     155 KCYEVGMVKGGSRKERCGYRLVRRQRSADE-QL-HCAGKAKRSGGHAPRVR-ELLLDA-L 210

Db     684 ip-plisllmslepdvlyaghdstkpdtssslltslnqlgerqllsvvkwskslpgfrnl 742
            |  |:  |:  ||   :  :: :: | | :|:: ||  |::::|: :: |:| :||| :|
Qy     211 SPEQLVLTLLEAEPPHVLISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVEL 269

Db     743 hiddqitliqyswmslmvfglqwrsykhvsqqmlyfapdlilneqrmk-essfyslcltm 801
            : ||: |::  || ::: || |||  | :| :: |||||:|: :  |    :: :  :
Qy     270 SLFDQVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDML 327
```

```
Db     802 wqipqefvklqvsqeeflcmkvllllntipleglrsqtqfeemrssyirelikaiq- 857
               |  | : : |:||:| ::|||: :: | : ||  :  |  : :   :  |:
Qy     328 LATISRFRELKLQHKEYLCVKAMILLNS-SMYPLVTATQDADSSRKLAHLLNAVTDALVW 386

Db     858 lrqkgvvsssqrfyqltklldnlhdlvkq 886
            :  |: :|| |:  :|::||  |   |::
Qy     387 VIAKSGISSQQQSMRLANLL-MLLSHVRH 414


RESULT   15
ENTRY            A25923     #type complete
TITLE            progesterone receptor - rabbit
ORGANISM         #formal_name Oryctolagus cuniculus #common_name domestic
                   rabbit
DATE             15-Aug-1988 #sequence_revision 15-Aug-1988 #text_change
                   03-Jun-1996
ACCESSIONS       A25923
REFERENCE        A25923
   #authors      Loosfelt, H.; Atger, M.; Misrahi, M.; Guiochon-Mantel, A.;
                   Meriel, C.; Logeat, F.; Benarous, R.; Milgrom, E.
   #journal      Proc. Natl. Acad. Sci. U.S.A. (1986) 83:9045-9049
   #title        Cloning and sequence analysis of rabbit progesterone receptor
                   complementary DNA.
   #cross references MUID:87067449
   #accession    A25923
      ##molecule_type mRNA
      ##residues       1-930 ##label LOO
CLASSIFICATION   #superfamily progesterone receptor, erbA transforming protein
                   homology
KEYWORDS         DNA binding; nucleus; steroid hormone receptor; zinc finger
FEATURE
   566-826             #domain erbA transforming protein homology #label ERBA\
   568-588             #region zinc finger\
   604-628             #region zinc finger
SUMMARY          #length 930  #molecular-weight 98666  #checksum 2858

  Query Match             21.2%;  Score 662;  DB 7;  Length 930;
  Best Local Similarity   32.5%;  Pred. No. 1.28e-97;
  Matches  106;  Conservative   91;  Mismatches 114;  Indels  15;  Gaps  15;

Db     567 iclicgdeasgchygvltcgsckvffkrameggqhnylcagrndcivdkirrkncpacrlr 626
           :| :|:| ||| |||| :| :|| ||||:::|:::|:|:: |:| :|| |||:| |||||
Qy      95 FCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLR 154

Db     627 kcqqagmvl-ggrkfk-kfnkvrvmraldavalpqpvgipnesq ritfspgqdiq lip 682
           || : |||  |:|: :   :  ||  |: |    | :  | :: | :               :: | |
Qy     155 KCYEVGMVKGGSRKERCGYRLVRRQRSADE-QL-HCAGKAKRSGGHAPRVRELLLDALSP 212

Db     683 -plinllmsiepdvlyaghdntkpdtssslltslnqlgerqllsvvkwskslpqfrnlhi 741
             :   :  ||        .      | |  :|:  ||  |     |       | | :||| :| :
Qy     213 EQLVLTLLEAEPPHVLISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSL 271

Db     742 ddqitliqyswmslmvfglgwrsykhvsqqmlyfapdlilneqrmk-essfyslcltmwq 800
            ||. |..  || ... || |||  | .| :: |||||:': :  |    ::  :   :
Qy     272 FDQVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDMLLA 329

Db     801 ipqefvklqvsqeeflcmkvllllnt-i-pleglrsqsqfeemrssyirelikaiq-lrq 857
                |  | : : |:||:| ::|||: : ||     :::   : :  :  : |: :
Qy     330 TTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIA 389

Db     858 kgvvsssqrfyqltklldnlhdlvkq 883
           |. :|| |:  :|::||  |   |::
Qy     390 KSGISSQQQSMRLANLL-MLLSHVRH 414


Search completed: Thu Dec 18 ... 1997
Job time : 103 secs.
```

```
*****************************************************************
   __   __   ______   _____   _____   _____   _   _
  |  \_/  | |  __  | |  ___| |  _  | |  ___| | | | |
  |       | | |__| | | |___  | |_| | | |     | |_| |
  | |\_/| | |  ____| |___  | |  _ <  | |     |  _  |
  | |   | | | |       ___| | | | \ \ | |___  | | | |
  |_|   |_| |_|      |_____| |_|  \_\|_____| |_| |_|  (TM)

*****************************************************************

          Release 2 1D John F  Collins  Biocomputing Research Unit
          Copyright (c) 1993, 1994, 1995  University of Edinburgh, U.K.
                  Distribution rights by IntelliGenetics, Inc

MPsrch_nn   n.a  -  n.a  database search, using Smith-Waterman algorithm

Run on:           Thu Dec 18 19:29:04 1997;  MasPar time 1063.37 Seconds
                                             1299.269 Million cell updates/sec
Tabular output not generated.

Title:            >US-08-826-361A-20
Description:      (1-1257) from US08826361A seq
Perfect Score     1257
N.A. Sequence:        1 ATGAATTACAGCATTCCCAG..........GGCATGCGAGGTCTGCCTGA 1257
     Comp:              TACTTAATGTCGTAAGGGTC          CCGTACGCTCCAGACGGACT

Scoring table:    TABLE default
                  Gap 6

Nmatch   STD :    Dbase 0;  Query 0

Searched:         352057 seqs, 549138275 bases x 2

Post-processing:  Minimum Match 0%
                  Listing first 45 summaries

Database:         embl-new3
                    1:BCT 2:FUN 3:GEN1 4:GEN2 5:HTG1 6:HTG2 7:HUM 8:INV
                    9:ORG 10:MAM 11:VRT 12:PLN 13:PRO 14:ROD 15:SYN 16:UNC
                    17:VIR
Database:         genbank99
                    18:BCT1 19:BCT2 20:BCT3 21:BCT4 22:BCT5 23:BCT6 24:BCT7
                    25:BCT8 26:BCT9 27:BCT10 28:BCT11 29:GEN1 30:GEN2
                    31:GEN3 32:HTG1 33:HTG2 34:HTG3 35:INV1 36:INV2 37:INV3
                    38:INV4 39:INV5 40:INV6 41:INV7 42:INV8 43:INV9 44:INV10
                    45:INV11 46:MAM1 47:MAM2 48:MAM3 49:VRT1 50:VRT2 51:VRT3
                    52:VRT4 53:PAT1 54:PAT2 55:PAT3 56:PAT4 57:PAT5 58:PHG
                    59:PLN1 60:PLN2 61:PLN3 62:PLN4 63:PLN5 64:PLN6 65:PLN7
                    66:PLN8 67:PLN9 68:PLN10 69:PLN11 70:PRI1 71:PRI2
                    72:PRI3 73:PRI4 74:PRI5 75:PRI6 76:PRI7 77:PRI8 78:PRI9
                    79:PRI10 80:PRI11 81:PRI12 82:PRI13 83:PRI14 84:PRI15
                    85:ROD1 86:ROD2 87:ROD3 88:ROD4 89:ROD5 90:ROD6 91:ROD7
                    92:ROD8 93:STS 94:SYN 95:UNA 96:VRL1 97:VRL2 98:VRL3
                    99:VRL4 100:VRL5 101:VRL6 102:VRL7 103:VRL8 104:VRL9
                    105:VRL10
Database:         genbank-new3
                    106:BCT 107:GEN1 108:GEN2 109:HTG1 110:HTG2 111:INV
                    112:MAM 113:VRT 114:PHG 115:PLN 116:PRI1 117:PRI2
                    118:ROD 119:SYN 120:UNA 121:VRL
Database:         u-embl50_99
                    122:part1

Statistics:       Mean 11.301;  Variance 4.407;  scale 2.565

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution

                              SUMMARIES
```

```
Result          Query
   No.  Score   Match Length DB  ID           Description             Pred. No.
-------------------------------------------------------------------------------
     1   1147    91.2   1560 76  HSRNAERB     H.sapiens mRNA for es   0.00e+00
     2    561    44.6   1599 91  RNU57439     Rattus norvegicus est   1.02e-00
     3    261    20.8    232 79  CJERB        C.jacchus mRNA for es   6.19e-236
     4    170    13.5   1728 51  ORZMER       Killifish mRNA for es   8.82e-140
     5    155    12.3   2085 50  GGERR        Chicken mRNA for estr   2.41e-124
     6    149    11.9   2792 52  ZFIER        Poephila guttata estr   5.45e-118
     7    148    11.8   1966 47  OAESTREC     O.aries mRNA for estr   5.87e-117
     8    140    11.1   1662 52  SSOESREC     S.salar mRNA for oest   1.01e-108
     9    140    11.1   1788 48  SSESTREC     S.scrofa mRNA for est   1.01e-108
    10    140    11.1   1880 91  RNESTROR     Rat mRNA for estrogen   1.01e-108
    11    140    11.1   2090 91  RNERR        Rat mRNA for estrogen   1.01e-108
    12    140    11.1   2090 92  S52129       estrogen receptor [ra   1.01e-108
    13    136    10.8   2106 78  HSU47678     Human 80 kDa estrogen   1.29e-104
    14    135    10.7   1279 52  SMOERA       Rainbow trout estroge   1.36e-103
    15    135    10.7   2092 55  I08528       Sequence 1 from Paten   1.36e-103
    16    135    10.7   2092 90  HUMERMCF     Human estrogen recept   1.36e-103
    17    135    10.7   4963 56  I15368       Sequence 16 from pate   1.36e-103
    18    135    10.7   6450 74  HSERR        Human mRNA for oestr    1.36e-103
    19    133    10.6   2619 90  HUMERVBCA    Human estrogen recept   1.52e-101
    20    132    10.5    923 51  S79923       estrogen receptor=rep   1.60e-100
    21    132    10.5   2078 89  MUSESTRU     Mouse estrogen recept   1.60e-100
    22    126    10.0   2443 51  XELESTRREC   Xenopus laevis estrog   2.10e-94
    23    126    10.0   2479 52  XELESTRC     Xenopus laevis estrog   2.10e-94
    24    126    10.0   2533 52  XELER3       Xenopus laevis estrog   2.10e-94
    25    113     9.0    586 79  HSZ75126     H sapiens mRNA for es   3.04e-81
    26    109     8.7    335 47  OAU30299     Ovis aries estrogen r   3.17e-77
    27     92     7.3    477 46  BTU64962     Bos taurus estrogen r   2.42e-60
    28     79     6.2    323 74  HSEST        H. sapiens mRNA for e   9.58e-47
    29     79     6.2    323 117  HSEST        H. sapiens mRNA for e   9.58e-47
    30     77     6.1    400 46  BTEREC       B taurus mRNA for est   8.70e-46
    31     74     5.9    365 51  OAESREXN4    O.aureus estrogen rec   6.32e-43
    32     71     5.6   3460 57  I25343       Sequence 1 from paten   4.39e-40
    33     68     5.4   2153 76  HSSTHOR2     Human mRNA for steroi   2.90e-37
    34     68     5.4   2153 55  I08706       Sequence 13 from Pate   2.90e-37
    35     66     5.3   1344 122  MMERR2GEN    M.musculus mRNA for e   2.14e-35
    36     66     5.3   1344 92  S82459       mERR-2=estrogen respo   2.14e-35
    37     64     5.1   2265 95  MMGRLA       Mouse mRNA for glucoc   1.53e-33
    38     63     5.0   1391 48  RAHPGR       Rabbit progesterone r   1.29e-32
    39     63     5.0   6244 56  I15367       Sequence 15 from pate   1.29e-32
    40     62     4.9   2376 10  SOU87953     Saguinus oedipus gluc   1.07e-31
    41     62     4.9   2376 116  SOU87953     Saguinus oedipus gluc   1.07e-31
    42     62     4.9   2376 10  ANU87952     Aotus nancymaae gluco   1.07e-31
    43     62     4.9   2376 116  ANU87952     Aotus nancymaae gluco   1.07e-31
    44     62     4.9   2575 95  MMGCRR       Mouse mRNA for glucoc   1.07e-31
    45     61     4.9   2641 48  TBGCR        T belangeri mRNA for    8.87e-31

                            ALIGNMENTS

RESULT     1
LOCUS       HSRNAERB     1560 bp    RNA             PRI       15-JUL-1996
DEFINITION  H.sapiens mRNA for estrogen receptor.
ACCESSION   X99101
NID         g1518262
KEYWORDS    estrogen receptor.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1560)
  AUTHORS   Mosselman,S., Polman,J. and Dijkema,R.
  TITLE     ER beta: identification and characterization of a novel human
            estrogen receptor
  JOURNAL   FEBS Lett. 392 (1), 49-53 (1996)
  MEDLINE   96354875
REFERENCE   2  (bases 1 to 1560)
  AUTHORS   Mosselman,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-JUL-1996) S. Mosselman, N.V.Organon, Biotechnology
            and Biochemistry, PO box 20 Oss, Molenstraat 110 Oss, 5340 BH,
            NETHERLANDS
FEATURES             Location/Qualifiers
     source          1..1560
                     /organism="Homo sapiens"
                     /tissue_type="testis"
                     /dev_stage="adult"
                     /clone_lib="Clontech  HL1010B"
                     /clone_lib="Clontech; 7414-1"
     CDS             19..1452
                     /codon_start=1
                     /product="estrogen receptor beta"
                     /db_xref="PID:e255390"
                     /db_xref="PID:g1518262"
                     /translation="MNYSIPSNVTNLEGGPGRQTTSPNVLWPTPGHLSPLVVHRQLSH
                     LYAEPQKSPWCEARSLEHTLPVNRETLKRKVSGNRCASPVTGPGSKRDAHFCAVCSDY
                     ASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRKCYEVG
                     MVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQLVLT
                     LLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFDQVR
                     LLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSRF
                     RELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKSG
                     ISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNAHVLRGC
                     KSSITGSECSPAEDSKSKEGSQNPQSQ"
BASE COUNT      367 a    408 c    445 g    340 t
ORIGIN

  Query Match             99.2%;  Score 1247;  DB 76;  Length 1560;
  Best Local Similarity  100.0%;  Pred. No. 0.00e+00;
  Matches  1247;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Db         19 atgaattacagcattcccagcaatgtcactaacttggaaggtgggcctggtcggcagacc 78
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy          1 ATGAATTACAGCATTCCCAGCAATGTCACTAACTTGGAAGGTGGGCCTGGTCGGCAGACC 60

Db         79 acaagcccaaatgtgttgtggccaacacctgggcacctttctcctttagtggtccatcgc 138
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy         61 ACAAGCCCAAATGTGTTGTGGCCAACACCTGGGCACCTTTCTCCTTTAGTGGTCCATCGC 120

Db        139 cagttatcacatctgtatgcggaacctcaaaagagtccctggtgtgaagcaagatcgcta 198
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        121 CAGTTATCACATCTGTATGCGGAACCTCAAAAGAGTCCCTGGTGTGAAGCAAGATCGCTA 180

Db        199 gaacacaccttacctgtaaacagagagacactgaaaaggaaggttagtgggaaccgttgc 258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        181 GAACACACCTTACCTGTAAACAGAGAGACACTGAAAAGGAAGGTTAGTGGGAACCGTTGC 240

Db        259 gccagccctgttactggtccaggttcaaagagggatgctcacttctgcgctgtctgcagc 318
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        241 GCCAGCCCTGTTACTGGTCCAGGTTCAAAGAGGGATGCTCACTTCTGCGCTGTCTGCAGC 300

Db        319 gattacgcatcgggatatcactatggagtctggtcgtgtgaaggatgtaaggcctttttt 378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        301 GATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATGTAAGGCCTTTTTT 360

Db        379 aaaagaagcattcaaggacataatgattatatttgtccagctacaaatcagtgtacaatc 438
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        361 AAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCAGTGTACAATC 420

Db        439 gataaaaaccggcgcaagagctgccaggcctgccgacttcggaagtgttacgaagtggga 498
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        421 GATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTACGAAGTGGGA 480

Db        499 atggtgaagtgtggctcccggagagagagatgtgggtaccgccttgtgcggagacagaga 558
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        481 ATGGTGAAGTGTGGCTCCCGGAGAGAGAGATGTGGGTACCGCCTTGTGCGGAGACAGAGA 540

Db        559 agtgccgacgagcagctgcactgtgccggcaaggccaagagaagtggcggccacgcgccc 618
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        541 AGTGCCGACGAGCAGCTGCACTGTGCCGGCAAGGCCAAGAGAAGTGGCGGCCACGCGCCC 600

Db        619 cgagtgcgggagctgctgctggacgccctgagccccgagcagctagtgctcaccctcctg 678
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        601 CGAGTGCGGGAGCTGCTGCTGGACGCCCTGAGCCCCGAGCAGCTAGTGCTCACCCTCCTG 660

Db        679 gaggctgagccgccccatgtgctgatcagccgccccagtgcgcccttcaccgaggcctcc 738
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        661 GAGGCTGAGCCGCCCCATGTGCTGATCAGCCGCCCCAGTGCGCCCTTCACCGAGGCCTCC 720

Db        739 atgatgatgtccctgaccaagttggccgacaaggagttggtacacatgatcagctgggcc 798
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        721 ATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTACACATGATCAGCTGGGCC 780

Db        799 aagaagattcccggctttgtggagctcagcctgttcgaccaagtgcggctcttggagagc 858
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        781 AAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAAGTGCGGCTCTTGGAGAGC 840

Db        859 tgttggatggaggtgttaatgatggggctgatgtggcgctcaattgaccaccccggcaag 918
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        841 TGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCAATTGACCACCCCGGCAAG 900

Db        919 ctcatctttgctccagatcttgttctggacagggatgaggggaaatgcgtagaaggaatt 978
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        901 CTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGGAAATGCGTAGAAGGAATT 960

Db        979 ctggaaatctttgacatgctcctggcaactacttcaaggtttcgagagttaaaactccaa 1038
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        961 CTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTTCGAGAGTTAAAACTCCAA 1020

Db       1039 cacaaagaatatctctgtgtcaaggccatgatcctgctcaattccagtatgtaccctctg 1098
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1021 CACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAATTCCAGTATGTACCCTCTG 1080

Db       1099 gtcacagcgacccaggatgctgacagcagccggaagctggctcacttgctgaacgccgtg 1158
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1081 GTCACAGCGACCCAGGATGCTGACAGCAGCCGGAAGCTGGCTCACTTGCTGAACGCCGTG 1140

Db       1159 accgatgctttggtttgggtgattgccaagagcggcatctcctcccagcagcaatccatg 1218
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1141 ACCGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATCTCCTCCCAGCAGCAATCCATG 1200

Db       1219 cgcctggctaacctcctgatgctcctgtcccacgtcaggcatgcgag 1265
              |||||||||||||||||||||||||||||||||||||||||||||||
Qy       1201 CGCCTGGCTAACCTCCTGATGCTCCTGTCCCACGTCAGGCATGCGAG 1247


RESULT      2
LOCUS       RNU57439     2555 bp    mRNA             ROD       12-JUN-1996
DEFINITION  Rattus norvegicus estrogen receptor beta mRNA, complete cds.
ACCESSION   U57439
NID         g1373280
KEYWORDS    .
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
            Murinae; Rattus.
REFERENCE   1  (bases 8 to 2555)
  AUTHORS   Kuiper,G.G., Enmark,E., Pelto-Huikko,M., Nilsson,S. and
            Gustafsson,J.A.
  TITLE     Cloning of a novel receptor expressed in rat prostate and ovary
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 93 (12), 5925-5930 (1996)
  MEDLINE   96234066
REFERENCE   2  (bases 1 to 2555)
  AUTHORS   Enmark,E.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-MAY-1996) Eva L. Enmark, Center for Biotechnology,
            NOVUM, Karolinska, Institute, NOVUM, Huddinge, S-141 57, Sweden
FEATURES             Location/Qualifiers
     source          1..2555
                     /organism="Rattus norvegicus"
                     /strain="Sprague Dawley"
                     /sex="male"
                     /tissue_type="prostate"
     CDS             417..1874
                     /note="nuclear receptor superfamily member"
                     /codon_start=1
                     /product="estrogen receptor beta"
```

```
                      /db_xref="PID:g1373291"
                      /translation="MTFYSPAVMNYSVPGSTSNLDGGPVRLSTSPNVLWPTSGHLSPL
                      ATHCQSSLLYAEPQKSPWCEARSLEHTLPVNRETLKRKLSGSSCASPVTSPNAKRDAH
                      FCPVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACR
                      LRKCYEVGMVKCGSRRERCGYRIVRRQRSSSEQVHCLSKAKRNGGHAPRVKELLLSIL
                      SPEQLVLTLLEAEPPNVLVSRPSMPFTEASMMMSLTKLADKELVHMIGWAKKIPGFVE
                      LSLLDQVRLLESCWMEVLMVGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDM
                      LLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLASANQEAESSRKLTHLLNAVTDAL
                      VWVIAKSGISSQQQSVRLANLLMLLSHVRHISNKGMEHLLSMKCKNVVPVYDLLLEML
                      NAHTLRGYKSSISGSECSSTEDSKNKESSQNLQSQ"
     misc_feature     726..923
                      /note="encodes DNA binding domain"
BASE COUNT      630 a    652 c    660 g    613 t
ORIGIN

  Query Match             68.5%;  Score 861;  DB 92;  Length 2555;
  Best Local Similarity   84.7%;  Pred. No. 0.00e+00;
  Matches 1051;  Conservative    0;  Mismatches  190;  Indels    0;  Gaps    0;

Db      441 atgaactacagtgttcccagcagcacagtaacctggacggtgggcctgtccgactgagc 500
            ||||| |||||  ||||| |||    || |||| |||| ||||||||||   || | || |
Qy        1 ATGAATTACAGCATTCCCAGCAATGTCACTAACTTGGAAGGTGGGCCTGGTCGGCAGACC 60

Db      501 acaagcccaaatgtgctatggccaacttctgggcacctgtctcctttagcgacccattgc 560
            ||||||||||||||| | ||||||||  |||||||||| |||||||||| |   |||| ||
Qy       61 ACAAGCCCAAATGTGTTGTGGCCAATACTTGGGCACCTTTCTCCTTTAGTGGTCCATCGC 120

Db      561 cagtcatcgctcctctatgcagaacctcaaagagtcccttggtgtgaagcaagatcacta 620
            || | ||| |   || ||||| |||||||||||||||||| |||||||||||||||| | |
Qy      121 CAGTTATCACATCTGTATGCGGAACCTCAAAAGAGTCCCTGGTGTGAAGCAAGATCGCTA 180

Db      621 gagcacaccttacctgtaaacagagagacactgaagaggaagcttagtgggagcagttgt 680
            || ||||||||||||||||||||||||||||||||| |||||| |||||||| | ||||
Qy      181 GAACACACCTTACCTGTAAATAGAGAGACACTGAAAAGGAAGGTTAGTGGGAACCGTTGC 240

Db      681 gccagccctgttactagtccaaacgcaaagagggatgctcacttctgccccgtctgcagc 740
            |||||||||||||||| |||||     |||||||||||||||||||||||| | |||||||
Qy      241 GCCAGCCCTGTTACTGGTCCAGGTTCAAAGAGGGATGTTCACTTCTGCCGTGTCTGCAGC 300

Db      741 gattatgcatctgggtatcattacggcgtttggtcatgtgaaggatgtaaggcctttttt 800
            ||||| ||||| || |||||| || || || ||||| ||||||||||||||||||||||||
Qy      301 GATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATGTAAGGCCTTTTTT 360

Db      801 aaaagaagcattcaaggacataatgattatatttgtcccagcacgaatcagtgtacaatc 860
            ||||||||||||||||||||||||||||||||||||| |||||||| || ||||||||| ||
Qy      361 AAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCAGTGTACAATC 420

Db      861 gacaggaacagcagcaagagctgccaggcttgccgacttcgcaagtgttatgaagtagga 920
            || || |||||||| || |||||||||||||||||||||||||| ||||||| ||||| |||
Qy      421 GATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTACGAAGTGGGA 480

Db      921 atggtcaagtgtggatccaggagagaacggtgtgggtaccgtatagtgcggaggcagaga 980
            ||||| ||||||||| ||| |||||||   | ||||||||||||   | ||||||||| ||||||
Qy      481 ATGGTGAAGTGTGGCTCCCGGAGAGAGAGATGTGGGTACCGCCTTGTGCGGAGACAGAGA 540

Db      981 agttctagccgagcaggtacactgcctgaggaaggcaaggaaacggtgggcatgcacccc 1040
            ||| |   ||||||| | |||||     |||| ||||||||||  || || || || |||
Qy      541 AGTGCCGACGAGCAGCTGCACTGTCCGGGCAAGGCCAAGAGAAGTGGCGGCCACGCGCCC 600

Db     1041 cgggtgaaggagctactgctgagcaccttgagtccagagcaactggtgctcaccctcctg 1100
            || |||  |||||| ||||||  | || |||| || ||||| || |||||||||||||||
Qy      601 CGAGTGCGGGAGCTGCTGCTGGACGCCCTGAGCCCCGAGCAGCTAGTGCTCACCCTCCTG 660

Db     1101 gaagctgaaccaccccatgtgctggtgagccgtcccagcatgcccttcaccgaggcctcc 1160
            || |||| || ||| |||||||| | |||| |||||    ||||||||||||||||||||
Qy      661 GAGGCTGAGCCGCCCCATGTGCTGATCAGCCGCCCCAGTGCGCCCTTCACCGAGGCCTCC 720

Db     1161 atgatgatgtccctcactaagctggcggacaaggaactggtgcacatgattggctgggcc 1220
            |||||||||||||| || ||| |||| ||||||||   |||| ||||||||   ||||||||
Qy      721 ATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTACACATGATCAGCTGGGCC 780

Db     1221 aagaaattccctggttttgtggagctcagcctatttgaccaagtccggctcttagaaagc 1280
            ||||| || || |||||||||||||||||||| ||||||||| |||||||| ||| |||
Qy      781 AAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAAGTGCGGCTCTTGGAGAGC 840

Db     1281 tgctggatggaggtgcttatggtggggctgatgtggcgctccattgaccaccccggcaag 1340
            || |||||||||||| ||||| |||| |||||||||||||| || |||||||||||||||
Qy      841 TGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCAATTGACCACCCCGGCAAG 900

Db     1341 ctcatttcgctcccgacctcgtttctggacagggatgaggggaagtgcgtagaagggatt 1400
            ||||| || ||||| || || |||||||||||||||||||||||| |||||||||| |||
Qy      901 CTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGGAAATGCGTAGAAGGAATT 960

Db     1401 ctggaaatctttgacatgctcctggcgacgacgtcaaggttccgtgagttaaaactccag 1460
            |||||||||||||||||||||||||||| || || |||||||| || |||||||||||
Qy      961 CTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTTCGAGAGTTAAAACTCCAA 1020

Db     1461 cacaaggagtatctctgtgtgaaggccatgatcctcctcaactccagtatgtaccccttg 1520
            ||||| || ||||||||||| |||||||||||||| ||||| |||||||||||||| ||
Qy     1021 CACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAATTCCAGTATGTACCCTCTG 1080

Db     1521 gcttctgcaaaccaggaggcagaaagtagccggaagctgacacacctactgaacgcggtg 1580
            |    | || | |||||| || || || |||||||||||| | ||| | ||||||| |||
Qy     1081 GTCACAGCCACCCAGGATGCTGACAGCAGCCGGAAGCTGGCTCACTTGCTGAACGCCGTG 1140

Db     1581 acagatgccctggtctgggtgattgcgaagagtggtatctcctcccagcagcagtcagtc 1640
            || |||||   |||| ||||||||||| ||||| || ||||||||||||||||| ||   |
Qy     1141 ACCGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATCTCCTCCCAGCAGCAATCCATG 1200

Db     1641 cgactggccaacctcctgatgcttctttctcacgtcaggca 1681
            || ||||| |||||||||||||| || || ||||||||||||
Qy     1201 CGCCTGGCTAACCTCCTGATGCTCCTGTCCCACGTCAGGCA 1241


RESULT      3
LOCUS       CJERB          332 bp    RNA             PRI       14-NOV-1996
DEFINITION  C.jacchus mRNA for estrogen receptor beta.
ACCESSION   Y09372
NID         g1669661
KEYWORDS    DNA binding; estrogen receptor; steroid hormone receptor;
            transactivation.
SOURCE      common marmoset.
  ORGANISM  Callithrix jacchus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Platyrrhini; Callitrichidae;
            Callithrix.
REFERENCE   1  (bases 1 to 332)
  AUTHORS   Gaughan,J.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 332)
  AUTHORS   Gaughan,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-NOV-1996) J. Gaughan, Medical Research Council,
            Centre for Reproductive Biology, 37 Chalmers Street, Edinburgh, EH3
            9EW, UK
COMMENT     Related entries: X99101 & U57439
            Related entries: X99101 & U57439.
FEATURES             Location/Qualifiers
     source          1..332
                     /organism="Callithrix jacchus"
                     /tissue_type="testis"
                     /tissue_type="ovary"
                     /tissue_type="prostate"
                     /clone="pGEM-T vector"
     CDS             18..>332
                     /function="steroid hormone receptor"
                     /codon_start=1
                     /product="estrogen receptor beta"
                     /db_xref="PID:e281483"
                     /db_xref="PID:g1669662"
                     /translation="MNYSIPSSVINLEAGPSRQIISFNMLWSTPDHLSFLAVHHQLSH
                     LYAEPQKTPWCFARSLEHTLPVSRETLKRKVSGNHCASPVIGPSSKRDAHFCAVCSDY
                     AAE"
BASE COUNT       91 a     84 c     77 g     80 t
ORIGIN
```

```
  Query Match             20.8%;  Score 261;  DB 70;  Length 332;
  Best Local Similarity   92.2%;  Pred. No. 6.19e-236;
  Matches  285;  Conservative    0;  Mismatches   24;  Indels    0;  Gaps    0;

Db       18 atgaattacagcattcccagcagtgttactaacttggaagcggggcctggtcggcagatc 77
            ||||||||||||||||||||||| ||| |||||||||||||  ||||||||||||||| |
Qy        1 ATGAATTACAGCATTCCCAGCAATGTCACTAACTTGGAAGGTGGGCCTGGTCGGCAGACC 60

Db       78 acaagcccaaatatgctgtggtcaacacctggcacctttctcctttagtgtccatcac 137
            ||||||||| ||| ||  ||||| ||||||||||| |||||||||||||||  ||||||| |
Qy       61 ATAAGCCCAAATGTGTTGTGGCCAACACCTGGGCACTTTCTCCTTTAGTGGTCCATCGC 120

Db      138 cagttatcacatctgtatgcagaacctcaaaagactccctggtgtgaagcaagatcacta 197
            |||||||||||||||||||| ||||||||||||| |||||||||||||||||||||| |||
Qy      121 CAGTTATCACATCTGTATGCGGAACCTCAAAAGACTCCCTGGTGTGAAGCAAGATCGCTA 180

Db      198 gaacataccttacctgtaagcagagagacactgaaaaggaaagttagtgggaaccattgt 257
            ||||| |||||||||||| ||||||||||||||||||||||| ||||||||||||| |||
Qy      181 GAACACACCTTACCTGTAAACAGAGAGACACTGAAAAGGAAGGTTAGTGGGAACCCTTGC 240

Db      258 gccagccctgttactggtccaagttcaaagagggatgctcacttctgtgcagtctgcagc 317
            |||||||||||||||||| |||||||||||||||||||||||||||| || |||||||||
Qy      241 GCCAGCCCTGTTACTGGTTCAGGTTCAAAGAGGGATGCTCACTTCTGCCCTGTCTGCAGC 300

Db      318 gattatgca 326
            ||||| |||
Qy      301 GATTACGCA 309


RESULT     4
LOCUS       ORZMER       1728 bp    mRNA              VRT       22-DEC-1994
DEFINITION  Killifish mRNA for estrogen receptor, complete cds.
ACCESSION   D28954
NID         g473547
KEYWORDS    estrogen receptor.
SOURCE      Oryzias sp. (strain d rR) (library: lambda ZAP II) liver cDNA to
            mRNA.
  ORGANISM  Oryzias sp.
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Actinopterygii; Neopterygii; Teleostei; Euteleostei;
            Acanthopterygii; Atherinomorpha; Cyprinodontiformes;
            Adrianichthyoidei; Oryziidae; Oryzias
REFERENCE   1  (bases 1 to 1728)
  AUTHORS   Okada,H
  JOURNAL   Unpublished (1994)
COMMENT     Submitted (15-Mar-1994) to DDBJ by:
            Hisashi Okada
            Center for Gene Science
            Hiroshima University
            1-4-2 Kagamiyama, Higashi Hiroshima
            Hiroshima 724
            Japan
            Phone: 0824-24-6283
            Fax:   0824-22-7184.
FEATURES             Location/Qualifiers
     source          1..1728
                     /organism="Oryzias sp."
                     /strain="d-rR"
                     /tissue_type="liver"
                     /clone_lib="lambda ZAP II"
     CDS             1..1728
                     /gene="MER"
                     /codon_start=1
                     /product="estrogen receptor"
                     /db_xref="PID:g602326"
                     /translation="MYPEESRGSGGVLLWIEWKGRTIMPAPNPAITPLYSQSSIGYYS
                     APLETNGPPSEGSLQSLGSGPTSPLVFVPSSPRLSPFMHPPSHHYLETTSTPVYRSSH
                     QGASREDQCGSRFDTCSLGELGAGAGAGGFEMAKDTRFCAVCSDYASGYHYGVWSCEG
                     CKAFFKRSIQGHNDYMCPATNQCTIDRNRRKSCQACRLRKCYEVGMMKGGVRKDRIRI
                     LRRDKRRTGVGDGDKVVKGQEHKTVHYDGRKRSSTGGGGGGGGGRLSVTSIPPEQVLL
                     LLQGAEPPILCSRQKLSRPYTEVTMMTLLTSMADKELVHMIAWAKKLPGFLQLSLHDQ
                     VLLLESSWLEVLMIGLIWRSIHCPGKLIFAQDLILDRNEGDCVEGMTEIFDMLLATAS
                     RFRVLKLKPEEFVCLKAIILLNSGAFSFCIGTMEPLHNSAAVQSMLDTITDALIHYIS
                     ISSYLADEQARSLACLLLILSHIRHMSNKGMEHLYSMKCKNKVPLYDLLLEMLDAHRL
                     HHPVRAPQSLSQVDRDPPSTSSGGGGIAPGSISASRGRIESPSRGPFAPSVLQYGGSR
                     PDCTPALQD"
BASE COUNT      379 a    513 c    496 g    340 t
ORIGIN

  Query Match             13.5%;  Score 170;  DB 51;  Length 1728;
  Best Local Similarity   73.2%;  Pred. No. 8.82e-140;
  Matches  268;  Conservative    0;  Mismatches   98;  Indels    0;  Gaps    0;

Db      883 ccgtacaccgaggtcaccatgatgaccctgctcaccagcatggcagacaaggagctggtc 942
            || | |||||||| | ||||||||      || ||||   |||| ||||||||| ||||
Qy      703 CCCTTCACCGAGGCCTCCATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTA 762

Db      943 cacatgatcgcctgggccaagaagctcccaggtttcctgcagctatccctgacagatcag 1002
            |||||||||  ||||||||||||| | || || ||| || ||||    ||||  ||| ||
Qy      763 CACATGATCAGCTGGGCCAAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAA 822

Db     1003 gtgctgctgctggagagctcgtggctggaggtgctcatgatcggcctcatttacaggtcc 1062
            |||| |||  |||||||||  ||| |||||||| | ||||| || || || ||| | ||
Qy      823 GTGCGGCTCTTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCA 882

Db     1063 atcaactgtcccgggaagctcatctttgcacaagacctcatcctggacaggaatgaggga 1122
            ||  ||   ||||| ||||||||||||| | ||| ||  | ||||||||| |||||||
Qy      883 ATTGACCACCCCGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGG 942

Db     1123 gactgcgtggaaggcatgacggagatcttcgacatgctgctggccactgcttcccgcttc 1182
             | ||||| ||||| ||   ||| ||||| |||||||| ||||| ||| ||||   | ||
Qy      943 AAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTT 1002

Db     1183 cgtgtgctcaaactcaacctgaggaattcgtctgcctcaaagctattatttactcaac 1242
            || | | | |||||| |||    | ||||    ||||  |||| || || ||   | ||||
Qy     1003 CGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAAT 1062

Db     1243 tccggt 1248
            ||| ||
Qy     1063 TCCAGT 1068


RESULT     5
LOCUS       GGERR        2085 bp    RNA               VRT       12-SEP-1993
DEFINITION  Chicken mRNA for estrogen receptor.
ACCESSION   X03805
NID         g63378
KEYWORDS    DNA binding protein; estrogen receptor; receptor; steroid hormone
            receptor; unidentified reading frame.
SOURCE      chicken.
  ORGANISM  Gallus gallus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Archosauria; Aves; Neognathae; Galliformes;
            Phasianidae; Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 2085)
  AUTHORS   Krust,A., Green,S., Argos,P., Kumar,V., Walter,P., Bornert,J.M. and
            Chambon,P.
  TITLE     The chicken oestrogen receptor sequence: homology with v-erbA and
            the human oestrogen and glucocorticoid receptors
  JOURNAL   EMBO J. 5 (5), 891-897 (1986)
  MEDLINE   86247578
COMMENT     Data kindly reviewed (23-JUL-1986) by S. Green.
FEATURES             Location/Qualifiers
     source          1..2085
                     /organism "Gallus gallus"
     mRNA            1..2085
                     /note="ER mRNA"
     CDS             69..131
                     /note="open reading frame 1 (aa 1-20)"
                     /codon_start=1
                     /db_xref="PID:g63379"
                     /translation="MYCFNMISAEPQNFFWCFF"
     CDS             177..200
                     /note="open reading frame 2 (aa 1-7)"
                     /codon_start=1
                     /db_xref="PID:g584490"
                     /translation="MFLAHCH"
     CDS             203..1972
                     /note="chicken estrogen receptor (aa 1-589)"
                     /codon_start=1
                     /db_xref="PID:g63380"
                     /db_xref="SWISS-PROT:P06212"
                     /translation="MIMTIHTKASGVTLLHQIQGIHLEILSRPQLKIPLERSLSDMYV
                     ESNKTGVFNYPEGATYDFGTTAPVYGSTTISYAPTSFSFGSSSLAGFHSLNNVPPSPV
                     VFLQTAPQLSPFIHHHSQQVPYYLENEPSGFGMREAAPPAFYRPSSTNRRHSIPERMS
                     STNEKGSLSMESTKETRYCAVCNDYASGYHYGVWSCEGCKAFFKRSIQGHNDYMCPAI
                     NQCTIDKNRRKSCQACRLRKCYEVGMMKGGIRKDRRGGEMMKQKRQREEQDSRNGFAS
                     STELRAPTLWTSPLVVKHNKKNSPALSLTAEQMVSALLEAEPPIVYSEYDPNRPFNEA
                     SMMTLLTNLADRELVHMINWAKRVPGFVDLTLHDQVHLLECAWLEILMIGLVWRSMEH
                     PGKLLFAPNLLLDRNQGKCVEGMVEIFDMLLATAARFRMMNLQGEEFVCLKSIILLNS
                     GVYTFLSSTLKSLEEKDYIHRVLDKITDTLIHLMAKSGLSLQQQHRRLAQLLLILSHI
                     RHMSNKGMEHLYNMKCKNVVPLYDLLLEMLDAHRLHAPAARSAAPMEEENRNQLTTAP
                     ASSHSLQSFYINSKEEESMQNTI"
     misc_feature    203..313
                     /note="aa 1-37 -> highly conserved region A"
     misc_feature    314..721
                     /note="conserved region B (aa 38-173)"
     misc_feature    722..970
                     /note="highly conserved region C (aa 174-256) put
                     DNA-binding domain"
     misc_feature    971..1087
                     /note="conserved region D (aa 257-295)"
     misc_feature    1088..1840
                     /note="highly conserved region E (aa 296-546, putative
                     estrogen binding domain"
     misc_feature    1841..1969
                     /note="conserved region F (aa 547-589)"
BASE COUNT      582 a    519 c    480 g    504 t
ORIGIN

  Query Match             12.3%,  Score 155,  DB 50,  Length 2085;
  Best Local Similarity   65.2%;  Pred. No. 3.41e-124;
  Matches   364;  Conservative    0;  Mismatches 191;  Indels    3;  Gaps    2;

Db   1180 ccccaatagaccatttaatgaagcatctatgatgaccctgttgaccaaccttgcagatag 1239
            ||||| |   || |||| ||| || || |||||||       ||||||| | || ||||
Qy    693 CCCCAGTGCGCCTTTCACCGAGGCCTCCATGATGATGTCCCTGACCAAGTTGGCCGACAA 752

Db   1240 agaattagtgcacatgattaactgggcaaagagagttccaggatttgtggatttaacact 1299
            || || || ||||||||| |||||| ||||   |||| || ||||||||  | |  ||
Qy    753 GGAGTTGGTACACATGATCAGCTGGGCCAAGAAGATTCCCGGCTTTGTGGAGCTCAGCCT 812

Db   1300 ccatgatcaggtccatctgctggaatgtgcctggttagagatattgatgattggtcttgt 1359
                || || || |  ||  ||||  |     ||| | ||| | || ||||| ||  |  |
Qy    813 GTTCGATCAAGTGCGGCTCTTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGAT 872

Db   1360 ctggcgctcaatggaacatccagggaagcttttatttgcacctaatctattactggacag 1419
            ||||||| || || ||||| || ||||| | ||||| ||  |||  | ||||||||
Qy    873 GTGGCGCTCAATTGACCACCCCGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAG 932

Db   1420 gaatcaaggaaaatgtgtagaaggcatggtggaaatctttgacatgctactggctactgc 1479
            | ||   |||||||| ||||| || ||  |||||||||||||||||||| ||||| ||| |
Qy    933 GGATGAGGGGAAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTAC 992

Db   1480 tgctcggtttcggatgatgaaccttcaaggagaggaatttgtgtgccttaagtccatcat 1539
            | |  |||||||    | | || || |||     | |||| | | ||  | ||| |||| ||
Qy    993 TTCAAGGTTTCGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGAT 1052

Db   1540 cctgctcaattctggtgtgtacacttttctttctagcaccttgaaatctctggaagagag 1599
            ||||||||||||   || ||||| || |   |   |     |||  | |  ||    |  ||
Qy   1053 CCTGCTCAATTCCAGTATGTACCCTCTGGTCACAGCGACCCAGGATGCTGAC-AGCAGCC 1111

Db   1600 ggactatatccaccgtgttctggacaaaatcacagatactctgatacacctaatggcaaa 1659
            |||   |   |    | ||  ||| ||      | || ||| || || |       | || | ||
Qy   1112 GGAAGCTGGCTCACTTG--CTGAACGCCGTGACCGATGCTTTGGTTTGGGTGATTGCCAA 1169

Db   1660 gtcaggtctttctctgcagcagcaacaccggcgactagctcagctcctccttatcctctc 1719
            |   ||  |  |    ||||||||| |  || || | | ||||| |  |||| ||
Qy   1170 GAGCGGCATCTCCTCCCAGCAGCAATCCATGCGCCTGGCTAACCTCCTGATGCTCCTGTC 1229

Db   1720 tcacatcaggcatatgag 1737
            ||| |||| ||||  |.|
Qy   1230 CCACGTCAGGCATGCGAG 1247


RESULT     6
LOCUS       ZFIER        2792 bp    mRNA             VRT       22-JUL-1996
DEFINITION  Poephila guttata estrogen receptor mRNA, complete cds.
ACCESSION   L79911
NID         g1449145
KEYWORDS    estrogen receptor.
SOURCE      Poephila guttata female adult hypothalamus-preoptic area cDNA to
            mRNA.
  ORGANISM  Poephila guttata
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Archosauria; Aves; Neognathae; Passeriformes;
            Estrildidae, Estrildinae; Poephila.
REFERENCE   1  (bases 1 to 2792)
  AUTHORS   Jacobs,E.C., Arnold,A.P. and Campagnoni,A.T.
  TITLE     Zebra finch estrogen receptor cDNA: cloning and mRNA expression
  JOURNAL   J. Steroid Biochem. Mol. Biol. 59(2) (1996) In press
FEATURES             Location/Qualifiers
     source          1..2792
                     /organism="Poephila guttata"
                     /dev_stage="adult"
                     /sex="female"
                     /tissue_type="hypothalamus-preoptic area"
     5'UTR           <1..132
     exon            <1..560
                     /note="Exon boundaries based on alignment to chicken ER"
                     /number=1
     misc_feature    127..237
                     /note="domain A"
     CDS             133..1896
                     /codon_start=1
                     /product="estrogen receptor"
                     /db_xref="PID:g1449146"
                     /translation="MILHIKISGVTLLHQIQGTELETLSRPQLKIPLERSLSDMYVET
                     NKTGVFNYPEGATYDFGTTAPVYSSTTLSYAPTSESFGSSSLAGFHSLNSVPPSPVVF
                     LQTAPHWSPFIHHHSQQVPYYLENDQGSFGMREAAPPAFYRPNSDNRRHSIRERMSSA
                     NEKGSLSMESIKETRYCAVCNDYASGYHYGVWSCEGCKAFFKRSIQGHNDYMCPATNQ
                     CTIDKNRRKSCQACRLRKCYEVGMMKGGIRKDRRGGRVMKQKRQREEQDSRNGEASST
                     ELRAPTLWASPLVVKHNKKNSPALSLTAEQMVSALLEAEPPLVYSEYDPNRPFNEASM
                     MTLLTNLADRELVHMINWAKRVPGFVDLTLHDQVHLLECAWLEILMIGLVWRSMEHPG
                     KLLFAPNLLLDRNQGKCVEGMVEIFDMLLATAARFRMMNLQGEEFVCLKSIILLNSGV
                     YTFLSSTLKSLEEKDYIHRVLDKITDTLIHLMAKSGLSLQQQHRRLAQLLLILSHIRH
                     MSNKGMEHLYNMKCKNVVPLYDLLLEMLDAHRLHAPAARSAAPMEEENRSQLTTASAS
                     SHSLQSFYINSKEEENMQNTL"
     misc_feature    238..645
                     /note="domain B"
     exon            561..751
                     /number=2
     misc_feature    646..894
                     /note="domain C, DNA binding domain"
     exon            752..868
                     /number=3
     exon            869..1204
                     /number=4
     misc_feature    895..1011
                     /note="domain D"
     misc_feature    1012..1764
                     /note="domain E, ligand binding domain"
     exon            1205..1343
                     /number=5
     exon            1344..1477
                     /number=6
     exon            1478..1661
                     /number=7
     exon            1662..>2792
                     /number=8
     misc_feature    1765..1893
                     /note="domain F"
     3'UTR           1897..>2792
     polyA_signal    2737..2743
BASE COUNT      794 a    669 c    615 g    714 t
ORIGIN

  Query Match             11.9%;  Score 149;  DB 52;  Length 2792;
  Best Local Similarity   69.9%;  Pred. No. 5.45e-118;
  Matches  262;  Conservative    0;  Mismatches 113;  Indels    0;  Gaps    0;

Db    1114 ccattcaatgaagcttctatgatgaccctgttgaccaaccttgcagacagagaattggtg 1173
           || ||||   || || || |||||||        ||||||   | || ||||  || |||||
Qy     703 CCCTTCACCGAGGCCTCCATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTA 762

Db    1174 cacatgatcaactgggcaaagagagttccaggatttgtggatttaacactccatgatcag 1233
           |||||||||| |||||| ||||     |||| || ||||||||   | |  ||     || ||
Qy     763 CACATGATCAGCTGGGCCAAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAA 822

Db    1234 gtccatctactggaatgtgcatggttagagatactgatgatcggcttagtgtggcgctcc 1293
           || |  ||  ||||  |    ||| | ||| |   | ||||| ||   |  ||||||||||
Qy     823 GTGCGGCTCTTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGGCCTCA 882

Db    1294 atggaacacccaggaaagcttttgtttgcacctaacctgttactggacaggaatcaaggg 1353
           || || ||||| || |||||   | ||||| ||   | ||   | ||||||||| || | |||
Qy     883 ATTGACCACCCCGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGG 942

Db    1354 aaatgtgtagagggcatggtggaaatctttgacatgctgctggctactgctgctcggttt 1413
           ||||| ||||| || ||  |||||||||||||||||| ||||| ||| || |  |||||
Qy     943 AAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTT 1002

Db    1414 cgcatgatgaaccttcaaggggaggaatttgtgtgccttaagtccatcatcctgctcaat 1473
           ||     | | || || |||      | |||| | | ||   | ||| |||| ||||||||||||
Qy    1003 CGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAAT 1062

Db    1474 tctggtgtgtacact 1488
           ||  || ||||| ||
Qy    1063 TCCAGTATGTACCCT 1077


RESULT     7
LOCUS       OAESTREC     1966 bp    RNA             MAM       08-OCT-1996
DEFINITION  O.aries mRNA for estrogen receptor.
ACCESSION   Z49257
NID         g1617201
KEYWORDS    estrogen receptor.
SOURCE      sheep.
  ORGANISM  Ovis aries
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Artiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Caprinae; Ovis.
REFERENCE   1  (bases 1 to 1966)
  AUTHORS   Madigou,T., Tiffoche,C., Lazennec,G., Pelletier,J. and
            Thieulant,M.L.
  TITLE     The sheep estrogen receptor: cloning and regulation of expression
            in the hypothalamo pituitary axis
  JOURNAL   Mol. Cell. Endocrinol. 121 (2), 153-163 (1996)
  MEDLINE   97047396
REFERENCE   2  (bases 1 to 1966)
  AUTHORS   Madigou,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-MAY-1995) Thierry Madigou, Interactions Cellulaires
            et Moleculaires, CNRS URA 256, Rennes, 35042, FRANCE
FEATURES             Location/Qualifiers
     source          1..1966
                     /organism="Ovis aries"
                     /strain="Romanov"
                     /clone="lambda P112"
                     /dev_stage="adult"
                     /tissue_type="anterior pituitary gland"
                     /clone_lib="cDNA library"
                     /sex="Female"
     mRNA            1..1966
                     /product="estrogen receptor"
BASE COUNT      439 a    604 c    549 g    374 t
ORIGIN

  Query Match             11.8%;  Score 148;  DB 47;  Length 1966;
  Best Local Similarity   69.9%;  Pred. No. 5.87e-117;
  Matches  260;  Conservative    0;  Mismatches 112;  Indels    0;  Gaps    0;

Db    1158 cctttcagtgaggcttcaatgatggccttgctgaccagcttacagacagagaagctaata 1217
           || ||||  ||||| || ||||||    |   ||||||     | || |||| |||| |||||
Qy     703 CCCTTCACCGAGGCCTCCATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTA 762

Db    1218 cacatgatcaactgggccaagagggtgccaggatttgtggatctggccctccacgatcaa 1277
           |||||||||| |||||| |||| | | || || |||||||| ||    |||    ||| |||
Qy     763 CACATGATCAGCTGGGCCAAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAA 822

Db    1278 gtccaccttttggaatgtgcctggctagagatcctcatgattggtcttgtctggcgctcc 1337
           || |  || |||||  |    ||| | ||| |   | ||||| || ||   | ||||||||
Qy     823 GTGCGGCTCTTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGGCCTCA 882

Db    1338 atggagcatccagggaagctcctatttgctcccaaccttctcctggacaggaaccaggga 1397
           || || || || || |||||| | ||||||||   | |||  | ||||||||| |   ||||
Qy     883 ATTGACCACCCCGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGG 942

Db    1398 aaatgtgtggaaggcatggtggagatctttgacatgttgctggctacatcgtctcggttc 1457
           ||||| || ||||| ||  |||| |||||||||||| | ||||| ||   | ||  ||||
Qy     943 AAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTT 1002

Db    1458 cgtatgatgaatctccagggagaagagtttgtgtgcctcaaatccatcatcttgcttaat 1517
           ||    | | || |||||       |||| | | | ||  ||||  |||| ||| |||| |||
Qy    1003 CGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAAT 1062

Db    1518 tctggggtgtac 1529
           ||  |  |||||
Qy    1063 TCCAGTATGTAC 1074


RESULT     8
LOCUS       SSOESREC     1662 bp    RNA             VRT       01-AUG-1995
DEFINITION  S.salar mRNA for oestrogen receptor.
ACCESSION   X89959
NID         g929901
KEYWORDS    oestrogen receptor.
SOURCE      Atlantic salmon.
  ORGANISM  Salmo salar
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Actinopterygii; Neopterygii; Teleostei; Euteleostei;
            Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
REFERENCE   1  (bases 1 to 1662)
  AUTHORS   Rogers,S.A., Llewellyn,L., Ramsurn,V.P., Sweeney,G.E. and Wigham,T.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1662)
  AUTHORS   Rogers,S.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JUL-1995) S.A. Rogers, School of Molecular and
            Medical Biosciences, University of Wales, PO Box 911, Cardiff, CF1
            3US, UK
FEATURES             Location/Qualifiers
     source          1..1662
                     /organism="Salmo salar"
                     /dev_stage="adult"
                     /tissue_type="liver"
     CDS             <1..1608
                     /codon_start=1
                     /product="oestrogen receptor"
                     /db_xref="PID:g929902"
                     /translation="SRMLTDPPRIGSMQSIGSSPTGPLVFVSSSPQLSPFLHPPGHHG
                     LPSQSYYLETSSTPLYRSSVVTNQLSASEEKLCITSNRQQSYAAAGSGVRVFEMANET
                     RYCAVCSDFASGYHYGFWSCEGCKAFFKRSIQGHNDYMCPATNQCTMDRNRRKSCQAC
                     RLRKCYEVGMVKGGLRKDRGGRVLRKDKRYCGPAGDREKPYGDLEHRTAPPQDGGRNS
                     SSSSLSGGGGWCGPRITMPPEQVLFLLQGAEPPALCSRQKVARPYTEVTMMTLLTSMA
                     DKELVHMIAWAKKVPGFQELSLHDQVQLLESSWLEVLMIGLIWRSIHCPGKLIFAQDL
                     ILDRSEGDCVEGMAEIFDMLLATVSRFRMLKLKREEFVCLKAIILLNSGAFSFCSNSV
                     ESLHNSSAVESMLDNITDALIHHISHSGASVQQQPRRQVQLLLLLSHIRHMSNKGMEH
                     LYSIKCKNKVPLYDLLLEMLDGHRLQSPGKVAQAGEQTEGPSTTTTTSTGSSIGPMRG
                     SQPTHIRSPGSGVLQYGSPSSDQMPIP"
BASE COUNT      383 a    481 c    461 g    336 t      1 others
ORIGIN

  Query Match             11.1%;  Score 140;  DB 52;  Length 1662;
  Best Local Similarity  69.3%;  Pred. No. 1.01e-108;
  Matches  251;  Conservative    0;  Mismatches 111;  Indels    0;  Gaps    0;

Db     784 ccctacaacaggtcaacatgatgacctctgctcaccagcatagctgacaaggagctgatg 843
           |||| ||| |||| | |||||||||       || ||||    ||||||||||||||| ||||
Qy     703 CCCTTCACCGAGGTTCTATGATGATGCTCCCTGACTAAGTTGGCCGACAAGGAGTTGGTA 762

Db     844 cacatgatcaacttggcctaagaaagtacctagatttcccgggactgtctctcatgacag 903
           |||||||||    ||||| |||||  | || || ||    ||||||     ||     |||||
Qy     763 CACATGATCAGCTGGGCCAAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAA 822

Db     904 gtgcagctgctggagattcctgactgaagtactgatgatcggactcatcatggaggtcc 963
           | || |||  ||||||| |  ||| |||||||| | ||||| || || || ||| | ||
Qy     823 GTGCGGCTCTTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCA 882

Db     964 atccactgccctggggaactcatctttgcccaggacctcatcctggacaggaatgaaggg 1023
           ||  ||  ||| || || |||||||||||| |  || ||  | |||||||||  ||| |||
Qy     883 ATTGACCACCCCGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGG 942

Db    1024 gactgtgtggaggtatggactgagatcttcgacatgctcctggccactgtgtctcgcttc 1083
            | || || || || ||     || ||||| ||||||||||||| |||    ||  | ||
Qy     943 AAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTT 1002

Db    1084 cgcatgctcaaactgaagcctgaggagtttgtgtgcctcaaagccatcatcttgctcaac 1143
           ||    | | |||||  | |   | || | | | | ||  |||| ||||| ||| |||||||
Qy    1003 CGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAAT 1062

Db    1144 tc 1145
           ||
Qy    1063 TC 1064


RESULT     9
LOCUS       SSESTREC     1788 bp    RNA             MAM       15-SEP-1994
DEFINITION  S.scrofa mRNA for estradiol receptor.
ACCESSION   Z37167
NID         g587554
KEYWORDS    estradiol receptor.
SOURCE      pig.
  ORGANISM  Sus scrofa
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Artiodactyla; Suiformes; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 1788)
  AUTHORS   Bokenkamp,D., Jungblut,P.W. and Thole,H.H.
  TITLE     The C-terminal half of the porcine estradiol receptor contains no
            post-translational modification: determination of the primary
            structure
  JOURNAL   Mol. Cell. Endocrinol. 104 (2), 163-172 (1994)
  MEDLINE   95080454
REFERENCE   2  (bases 1 to 1788)
  AUTHORS   Boekenkamp,D.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-SEP-1994) Dirk Boekenkamp, Max-Planck-Institut fuer
            experimentelle Endokrinologie, Feodor-Lynen Strasse 7, Hannover,
            30625, Germany
FEATURES             Location/Qualifiers
     source          1..1788
                     /organism="Sus scrofa"
                     /tissue_type="Uterus"
                     /sex="Female"
     CDS             1..1788
                     /codon_start=1
                     /product="estradiol receptor"
                     /db_xref="PID:g587555"
                     /translation="MTMTLHTKASGMALLHQIQANELEPLNRPQLKIPLERPLGEVYV
                     DSSKPAVYNYPEGAAYDFNAAAAASAPVYGQSGLAYGPGSEAAAFGANGLGGFQPLNS
                     VSPSPLVLLHPPPQLSPFLHPHGQQVPYYLENEPSGYAVREAGPPAFYRPNSDNRRQG
                     GRERLASTSDKGSMAMESAKETRYCAVCNDYASGYHYGVWSCEGCKAFFKRSIQGHND
```

```
                     YMCPATNQCTIDKNRRKSCQACRLRKCYEVGMMKGGIRKDRRGGRMLKHKRQRDDGEG
                     RNEAVPPGDMRSANLWPSPLLIKHTKKNSPVLSLTADQMISALLEAEPPIIYSEYDPT
                     RPLSEASMMGLLTNLADRELVHMINWAKRVPGFLDLSLHDQVHLLECAWLEILMIGLV
                     WRSMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEFVCLKS
                     IILLNSGVYTFLSSTLKSLEEKDHIHRVLDKIIDTLIHLMAKAGLTLQQQHRRLAQLL
                     LILSHIRHMSNKGMEHLYNMKCKNVVPLYDLLLEMLDAHRLHAPINLGGPPPEDMSQS
                     QLAISGSTPSHSLQMYYITGEAENFPTTI"
BASE COUNT      410 a    530 c    504 g    344 t
ORIGIN

  Query Match             11.1%;  Score 140;  DB 48;  Length 1788;
  Best Local Similarity  68.8%;  Pred. No. 1.01e-108;
  Matches  256;  Conservative    0;  Mismatches 116;  Indels    0;  Gaps    0;

Db    1006 cccctcagtgaggcttcaatgatgggcttgctgaccaacctcgcagacagggagctggta 1065
           ||| |||  ||||| || ||||||    |   ||||||||   | || |||| ||||  ||||
Qy     703 CCCTTCACCGAGGTTCTATGATGATGCTCCCTGACTAAGTTGGCCGACAAGGAGTTGGTA 762

Db    1066 cacatgatcaactgggcaaagagggtgccaggattttggatttaagcctccatgatcaa 1125
           |||||||||| |||||| |||| | | || || ||| ||||  | |||||     || |||
Qy     763 CACATGATCAGCTGGGCCAAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAA 822

Db    1126 gtgcatcttctggaatgtgcctggctagagatcctcatgattggtcttgtctggcgctcc 1185
           ||||  ||  ||||  |     ||| | |||  |   | ||||| || ||   ||||||||
Qy     823 GTGCGGCTCTTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCA 882

Db    1186 atggagcacccagggaagctcctgtttgctcctaacttgctcctggacaggaaccaggc 1245
           || || ||||| || ||||| | ||||||||  |   |   | |||||||||| |   ||||
Qy     883 ATTGACCACCCCGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGG 942

Db    1246 aagtgtgtcgagggaatggtggagatctttgacatgttgctggctacatcatctcgcttc 1305
           || || || || ||||| |||| ||||||||||| | ||||| ||  | ||  | ||
Qy     943 AAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTT 1002

Db    1306 cgtatgatgaatctccagggagaggagtttgtgtgcctcaaatccatcatttgcttaat 1365
           ||   | | || |||||      | || | | | || |||| |||| ||  |||| |||
Qy    1003 CGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAAT 1062

Db    1366 tctggagtgtac 1377
           ||  |  |||||
Qy    1063 TCCAGTATGTAC 1074


RESULT    10
LOCUS       RNESTROR     1880 bp    RNA             ROD       24-OCT-1996
DEFINITION  Rat mRNA for estrogen receptor.
ACCESSION   X61098
NID         g56120
KEYWORDS    estrogen receptor; steroid hormone receptor.
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 1880)
  AUTHORS   Maggi,A.M.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-JUN-1991) A.M.A. Maggi, Milano Mol Pharm Lab, Inst
            Pharmacological Sciences, Piazza Durante 11, 20131 Milano, ITALY
REFERENCE   2  (bases 1 to 1880)
  AUTHORS   Spreafico,E., Bettini,E., Pollio,G. and Maggi,A.
  TITLE     Nucleotide sequence of estrogen receptor cDNA from Sprague-Dawley
            rat
  JOURNAL   Eur. J. Pharmacol. 227 (3), 353-356 (1992)
  MEDLINE   93114361
FEATURES             Location/Qualifiers
     source          1..1880
                     /organism="Rattus norvegicus"
                     /strain="Sprague-Dawley"
                     /tissue_type="uterus"
     mRNA            1..1880
                     /evidence=experimental
                     /note="estrogen receptor"
     CDS             74..1876
                     /codon_start=1
                     /product="estrogen receptor"
                     /db_xref="PID:g56121"
                     /db_xref="SWISS-PROT:P06211"
                     /translation="MTMTLHTKASGMALLHQIQGNELEPLNRPQLKMPMERALGEVYV
                     DNSKPAVFNYPEGAAYEFNAAAAAAAAGASAPVYGQSSITYGPGSEAAAFGANSLGAF
                     PQLNSVSPSPLMLLHPPPHVSPFLHPHGHQVPYYLENEPSAYAVRDTGPPAFYRSNSD
                     NRRQNGRERLSSSSEKGNMIMESAKETRYCAVCNDYASGYHYGVWSCEGCKAFFKRSI
                     QGHNDYMCPATNQCTIDKNRRKSCQACRLRKCYEVGMMKGGIRKDRRGGRMLKHKRQR
                     DDLEGRNEMGTSGDMRAANLWPSPLVIKHTKKNSPALSLTADQMVSALLDAEPPLIYS
                     EYDPSRPFSEASMMGLLTNLADRELVHMINWAKRVPGFGDLNLHDQVHLLECAWLEIL
                     MIGLVWRSMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEF
                     VCLKSIILLNSGVYTFLSSTLKSLEEKDHIHRVLDKITDTLIHLMAKAGLTLQQQHRR
                     LAQLLLILSHIRHMSNKGMEHLYNMKCKNVVPLYDLLLEMLDAHRLHAPASRMGVPPE
                     EPSQSQLTTTSSTSAHSLQTYYIPPEAEGFPNTI"
BASE COUNT      438 a    548 c    505 g    389 t
ORIGIN

  Query Match             11.1%;  Score 140;  DB 91;  Length 1880;
  Best Local Similarity   82.4%;  Pred. No. 1.01e-108;
  Matches  178;  Conservative    0;  Mismatches   38;  Indels    0;  Gaps    0;

Db     628 ggagactcactactgtgctgtgtgtaacgactatgcctctggctaccattatggggtctg 687
           |||  ||| || ||| ||||| || | ||| || || || || || || ||||| |||||
Qy     273 GGATGCTCACTTCTGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTG 332

Db     688 gtcctgtgaaggctgcaaggctttctttaagagaagcattcaaggacataatgactacat 747
           ||| |||||||| || ||||| || ||||| ||||||||||||||||||||||| || ||
Qy     333 GTCGTGTGAAGGATGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATAT 392

Db     748 gtgtccagctacaaaccaatgcaccatcgataagaaccggaggaagagttgccaggcctg 807
            |||||||||||||| || || |  |||||||| |||||| | ||||| |||||||||||
Qy     393 TTGTCCAGCTACAAATCAGTGTATAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTG 452

Db     808 ccggctgcgcaagtgttacgaagtgggcatgatgaa 843
           ||| || || ||||||||||||||||| ||| ||||
Qy     453 CCGACTTCGGAAGTGTTACGAAGTGGGAATGGTGAA 488


RESULT     11
LOCUS       RNERR        2090 bp    RNA             ROD       12-SEP-1993
DEFINITION  Rat mRNA for estrogen receptor.
ACCESSION   Y00102
NID         g56110
KEYWORDS    estrogen receptor.
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 2090)
  AUTHORS   Muramatsu,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-MAR-1987) Masami Muramatsu,Dept. of Biochemistry,
            Faculty of Medicine The University of Tokyo, Hongo 7-3-1,
            Bunkyo-Ku, Tokyo 113 Japan
REFERENCE   2  (bases 1 to 2090)
  AUTHORS   Koike,S., Sakai,M. and Muramatsu,M.
  TITLE     Molecular cloning and characterization of rat estrogen receptor
            cDNA
  JOURNAL   Nucleic Acids Res. 15 (6), 2499-2513 (1987)
  MEDLINE   87174780
FEATURES             Location/Qualifiers
     source          1..2090
                     /organism="Rattus norvegicus"
                     /strain="Wistar"
                     /tissue_type="uterus"
                     /clone_lib="rat uterus cDNA 1."
     CDS             210..2012
                     /note="estrogen receptor protein (AA 1-600)"
                     /codon_start=1
                     /db_xref="PID:g56111"
                     /db_xref="SWISS-PROT:P06211"
                     /translation="MTMTLHTKASGMALLHQIQGNELEPLNRPQLKMPMERALGEVYV
                     DNSKPAVFNYPEGAAYEFNAAAAAAAAGASAPVYGQSSITYGPGSEAAAFGANSLGAF
                     PQLNSVSPSPLMLLHPPPHVSPFLHPHGHQVPYYLENEPSAYAVRDTGPPAFYRSNSD
                     NRRQNGRERLSSSSEKGNMIMESAKETRYCAVCNDYASGYHYGVWSCEGCKAFFKRSI
                     QGHNDYMCPATNQCTIDKNRRKSCQACRLRKCYEVGMMKGGIRKDRRGGRMLKHKRQR
                     DDLEGRNEMGTSGDMRAANLWPSPLVIKHTKKNSPALSLTADQMVSALLDAEPPLIYS
                     EYDPSRPFSEASMMGLLTNLADRELVHMINWAKRVPGFGDLNLHDQVHLLECAWLEIL
                     MIGLVWRSMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEF
                     VCLKSIILLNSGVYTFLSSTLKSLEEKDHIHRVLDKINDTLIHLMAKAGLTLQQQHRR
                     LAQLLLILSHIRHMSNKGMEHLYNMKCKNVVPLYDLLLEMLDAHRLHAPASRMGVPPE
                     EPSQSQLTTTSSTSAHSLQTYYIPPEAEGFPNTI"
BASE COUNT      488 a    619 c    537 g    446 t
ORIGIN

  Query Match             11.1%;  Score 140;  DB 91;  Length 2090;
  Best Local Similarity   82.4%;  Pred. No. 1.01e-108;
  Matches  178;  Conservative    0;  Mismatches   38;  Indels    0;  Gaps    0;

Db     764 ggagactcactactgtgctgtgtgtaacgactatgcctctggctaccattatggggtctg 823
           |||  ||| || ||| ||||| || | ||| || || || || || || ||||| |||||
Qy     273 GGATGCTCACTTCTGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTG 332

Db     824 gtcctgtgaaggctgcaaggctttctttaagagaagcattcaaggacataatgactacat 883
           ||| |||||||| || ||||| || ||||| ||||||||||||||||||||||| || ||
Qy     333 GTCGTGTGAAGGATGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATAT 392

Db     884 gtgtccagctacaaaccaatgcaccatcgataagaaccggaggaagagttgccaggcctg 943
            |||||||||||||| || || |  |||||||| |||||| | ||||| |||||||||||
Qy     393 TTGTCCAGCTACAAATCAGTGTATAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTG 452

Db     944 ccggctgcgcaagtgttacgaagtgggcatgatgaa 979
           ||| || || ||||||||||||||||| ||| ||||
Qy     453 CCGACTTCGGAAGTGTTACGAAGTGGGAATGGTGAA 488


RESULT     12
LOCUS       S52128       2090 bp    mRNA            ROD       12-MAR-1993
DEFINITION  estrogen receptor [rats, brain, mRNA, 2090 nt].
ACCESSION   S52128
NID         g262708
KEYWORDS    .
SOURCE      Rattus sp. brain.
  ORGANISM  Rattus sp.
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 2090)
  AUTHORS   Bettini,E., Pollio,G., Santagati,S. and Maggi,A.
  TITLE     Estrogen receptor in rat brain: presence in the hippocampal
            formation
  JOURNAL   Neuroendocrinology 56 (4), 502-508 (1992)
  MEDLINE   93116840
  REMARK    GenBank staff at the National Library of Medicine created this
            entry [NCBI gibbsq 121564] from the original journal article.
            This sequence comes from Fig. 1.
FEATURES             Location/Qualifiers
     source          1..2090
                     /organism="Rattus sp."
BASE COUNT      488 a    619 c    537 g    446 t
ORIGIN

  Query Match             11.1%;  Score 140;  DB 92;  Length 2090;
  Best Local Similarity   82.4%;  Pred. No. 1.01e-108;
  Matches  178;  Conservative    0;  Mismatches   38;  Indels    0;  Gaps    0;

Db     764 ggagactcactactgtgctgtgtgtaacgactatgcctctggctaccattatggggtctg 823
           |||  ||| || ||| ||||| || | ||| || || || || || || ||||| |||||
Qy     273 GGATGCTCACTTCTGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTG 332

Db     824 gtcctgtgaaggctgcaaggctttctttaagagaagcattcaaggacataatgactacat 883
           ||| |||||||| || ||||| || ||||| ||||||||||||||||||||||| || ||
Qy     333 GTCGTGTGAAGGATGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATAT 392
```

```
Db     884 gtgtccagctacaaatcagtgtacaatcgatgaagaaccgggggaggttgcccaggcctg 943
            |||||||||||||| || || || |||||||| |||||| | ||||| ||||||||||||
Qy     393 TTGTCCAGCTACAAATCAGTGTACAATCGATAAAAACCGGGGCAAGAGCTGCCAGGCCTG 452

Db     944 ccggctgcgcaagtgttacgaagtgggcatgatgaa 979
           ||| || || ||||||||||||||||| ||| ||||
Qy     453 CCGACTTCGGAAGTGTTACGAAGTGGGAATGGTGAA 488


RESULT    13
LOCUS       HSU47678     2106 bp    mRNA             PRI       17-MAY-1996
DEFINITION  Human 80 kDa estrogen receptor mRNA, containing exon 6 and 7
            duplication, complete cds.
ACCESSION   U47678
NID         g1197854
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2106)
  AUTHORS   Pink,J.J., Wu,S.Q., Wolf,D.M., Bilimoria,M.M. and Jordan,V.C.
  TITLE     A novel 80 kDa human estrogen receptor containing a duplication of
            exons 6 and 7
  JOURNAL   Nucleic Acids Res. 24 (5), 962-969 (1996)
  MEDLINE   96174665
REFERENCE   2  (bases 1 to 2106)
  AUTHORS   Jordan,C.V.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-JAN-1996) Craig V. Jordan, Robert H. Lurie Cancer
            Center, Northwestern University Medical Center, 303 E. Chicago
            Ave., Olson 8258, Chicago, IL 60611, USA
FEATURES             Location/Qualifiers
     source          1..2106
                     /organism="Homo sapiens"
                     /chromosome="6"
                     /map="6q25"
                     /cell_line="MCF-7"
                     /tissue_type="breast"
     exon            1..450
                     /number=1
     CDS             1..2106
                     /note="80 kDa estrogen receptor with an exon 6 and 7
                     repeat; It is able to bind estrogen"
                     /codon_start=1
                     /product="80 kDa estrogen receptor"
                     /db_xref="PID g1197855"
                     /translation="MTMTLHTKASGMALLHQIQGNELEPLNRPQLKIPLERPLGEVYL
                     DSSKPAVYNYPEGAAYEFNAAAAANAQVYGQTGLPYGPGSEAAAFGSNGLGGFPPLNS
                     VSPSPLMLLHPPPQLSPFLQPHGQQVPYYLENEPSGYTVREAGPPAFYRPNSDNRRQG
                     GRERLASTNDKGSMAMESAKETRYCAVCNDYASGYHYGVWSCEGCKAFFKRSIQGHND
                     YMCPATNQCTIDKNRRKSCQACRLRKCYEVGMMKGGIRKDRRGGRMLKHKRQRDDGEG
                     RGEVGSAGDMRAANLWPSPLMIKRSKKNSLALSLTADQMVSALLDAEPPILYSEYDPT
                     RPFSEASMMGLLTNLADRELVHMINWAKRVPGFVDLTLHDQVHLLECAWLEILMIGLV
                     WRSMEHPVKLLFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEFVCLKS
                     IILLNSGVYTFLSSTLKSLEEKDHIHRVLDKITDTLIHLMAKAGLTLQQQHQRLAQLL
                     LILSHIRHMRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEFVCLKSIILLNSGVYT
                     FLSSTLKSLEEKDHIHRVLDKITDTLIHLMAKAGLTLQQQHQRLAQLLLILSHIRHMS
                     NKGMEHLYSMKCKNVVPLYDLLLEMLDAHRLHAPTSRGGASVEETDQSHLATAGSTSS
                     HSLQKYYITGEAEGFPATV"
     exon            451..642
                     /number=2
     exon            643..759
                     /number=3
     exon            760..1095
                     /number=4
     exon            1096..1233
                     /number=5
     exon            1234..1368
                     /number=6
     exon            1369..1551
                     /number=7
     exon            1551..1685
                     /note="exon 6'"
     exon            1686..1868
                     /note="exon 7'"
     exon            1869..2106
                     /number=8
BASE COUNT      496 a    602 c    587 g    421 t
ORIGIN

  Query Match             10.8%;  Score 136;  DB 78;  Length 2106;
  Best Local Similarity   63.8%;  Pred. No. 1.29e-104;
  Matches  350;  Conservative    0;  Mismatches  190;  Indels    8;  Gaps    2;

Db    1006 ccttcagtgaagcttcgatgatgggcttactgaccaacttggcagacaggggactgatt 1065
           |||||||   || || || ||||||    |  ||||||||   |||| |||| || | ||||
Qy     703 CCCTTCAGCGAGGCCTCCATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTA 762

Db    1066 cacatgatcaactgggcgaagagggtgccaggctttgtggatttgaccctccatgatcag 1125
           ||||||||| ||||||  |||| |  | | || ||||||||    | | |||     || ||
Qy     763 CACATGATCAGCTGGGCCAAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTTGACCAA 822

Db    1126 gtccaccttctagaatgtgcctggctagagatcctgatgattggtctcgtctggcgctcc 1185
           || |  ||  | ||   |     ||| | |||  |  | ||||| || ||   | ||||||
Qy     823 GTGCGGCTCTTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCA 882

Db    1186 atggagcaccccgggaagctactgtttgctcctaacttgctcttggacaggaaccaggga 1245
           || || ||||| |  |||||  | ||||||||  |   |   |  |||||||| |   ||||
Qy     883 ATTGACCACCCCGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGG 942

Db    1246 aaatgtgtagagggcatggtggagatcttcgacatgctgctggctacatcatctcggttc 1305
           ||||| ||||| || ||  |||| ||||| ||||||||| ||||| ||  | ||  ||||
Qy     943 AAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTACCTCAAGGTTT 1002

Db    1306 cgcatgatgaatctgcagggagaggagtttgtgtgcctcaaatctattattttgcttaat 1365
           ||   | | || || ||     | || | | | | || |||| | || || |||  |||
Qy    1003 CGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAAT 1062

Db    1366 tctggagtgtacacatttctgtccagcaccctgaagtctctggaagagaaggaccatatc 1425
           ||  | ||||| |  |  |  |     |||| | | ||     |  || |||    |  |
Qy    1063 TCCAGTATGTACCCTCTGGTCACAGCGACCCAGGATGCTGAC-AGCAGCCGGAAGCTGGC 1121

Db    1426 caccgagtcctggacaagatcacagacactttgatccacctgatggccaaggcaggcctg 1485
              |  | ||| ||     | || || ||||| |      |||| ||||||    ||| |
Qy    1122 TCACTTG--CTGAACGCCGTGACCGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATC 1179

Db    1486 accctgcagcagcagcaccagcggctggcccagctcctcctcatcctctcccacatcagg 1545
            ||   ||||||||   |   ||| |||||  | |||||  |  |||| |||||| |||||
Qy    1180 TCCTCCCAGCAGCAATCCATGCGCCTGGCTAACCTCCTGATGCTCCTCTCCCACGTCAGG 1239

Db    1546 cacatgagg 1554
           ||   ||||
Qy    1240 CATGCGAGG 1248


RESULT    14
LOCUS       SMOERA       1278 bp    mRNA             VRT       15-JUN-1990
DEFINITION  Rainbow trout estrogen receptor mRNA, 3' end.
ACCESSION   M31559
NID         g213783
KEYWORDS    estrogen receptor.
SOURCE      Rainbow trout (female) liver, cDNA to mRNA.
  ORGANISM  Oncorhynchus mykiss
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Actinopterygii; Neopterygii; Teleostei; Euteleostei;
            Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
REFERENCE   1  (bases 1 to 1278)
  AUTHORS   Pakdel,F., Le Guellec,C., Vaillant,C., Le Roux,M.G. and
            Valotaire,Y.
  TITLE     Identification and estrogen induction of two estrogen receptors
            (ER) messenger ribonucleic acids in the rainbow trout liver:
            Sequence homology with other ERs
  JOURNAL   Mol. Endocrinol. 3, 44-51 (1989)
  MEDLINE   89127284
```

```
FEATURES             Location/Qualifiers
     source          1..1278
                     /organism="Oncorhynchus mykiss"
     CDS             <1..1278
                     /note="estrogen receptor"
                     /codon_start=1
                     /db_xref="PID g213784"
                     /translation="CSDFASGYHYGVWSCEGCKAFFKRSIQGHNDYMCPATNQCIMDP
                     NRRKSCQACRLRKCYEVGMVKGGLRKDRGGRVLRKLDKRYCGPAGDREKPIVTWSTGQR
                     PQDGGRNSSSSLNGGGGWRGPRITMPPEQVLFLLQGQTPALCSRQKVARPYTEVTMMT
                     LLTSMADKELVHMIAWAKKVPGFQELSLHDQVQLLESSWLEVLMIGLIWRSIHCPGKI
                     IFAQDLILDRSEGDCVEGMAEIFDMLLATVSRFGMLKLKPEEFVCLKAIILLNPGAFS
                     FCSNSVESLHNSSAVESMLDNITDALIHHISHSGASVQQQPRRQAQLLLLLSHIRHMS
                     NKGMEHLYSIKCKNKVPLYDLLLEMLDGHRLQAPGKVAQAGEQTEGPSTTTTTSTGSS
                     IGPMRGSQDTHIRSPGSGVLQYGSPSSDQMPIP"
BASE COUNT      294 a    369 c    366 g    249 t
ORIGIN

  Query Match             10.7%;  Score 135;  DB 52;  Length 1278;
  Best Local Similarity   68.8%;  Pred. No. 1.36e-103;
  Matches  247;  Conservative    0;  Mismatches  112;  Indels    0;  Gaps    0;

Db       454 ccctacacagaggtcaccatgatgaccctgctcaccagcatggctgacaaggagctggtg 513
             |||| ||| |||| | |||||||||       || ||||    |||| ||||||||| ||||
Qy       703 CCCTTCACCCAGCCCTCCATCATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTA 762

Db       514 cacatgatcgcttgggctaagaaagtaccaggtttccaggagctgtctctccatgaccag 573
             |||||||||     ||||| |||||   | || || ||    ||||||     ||     |||||
Qy       763 CACATGATCAGCTGGGCCAAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAA 822

Db       574 gtgcagctgctggagagttcctggctggaggtgctgatgatcggactcatatggcggtcc 633
             |||| |||   ||||||| |   ||| |||||||| | ||||| || || || ||||| ||
Qy       823 GTGCGGCTCTTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCA 882

Db       634 atccactgccctgggaaactcatcttcgcccaggacctcatactggacaggagtgaaggg 693
             ||   ||   ||| || || ||||||||| || |   || ||   | |||||||||   ||| |||
Qy       883 ATTGACCACCCCGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGG 942

Db       694 gactgtgtggagggtatggctgagatcttcgacatgctcctggccactgtgtctcgcttc 753
              | || || || || ||      || ||||| |||||||||||||| |||     ||   | ||
Qy       943 AAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTT 1002

Db       754 ggcatgcttaaactgaagcctgaggagtttgtgtgcctcaaagccatcatcttgctcaa 812
              |    | | |||||   | |     | || | | | | ||   |||| ||||| ||| |||||||
Qy      1003 CGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAA 1061


RESULT   15
LOCUS       I08538       2092 bp                  PAT       14-NOV-1994
DEFINITION  Sequence 1 from Patent WO 8705049.
ACCESSION   I08538
NID         g588753
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2092)
  AUTHORS   Shine,J.
  TITLE     EUCARYOTIC EXPRESSION OF STEROID RECEPTOR PROTEINS
  JOURNAL   Patent: WO 8705049-A 1 27-AUG-1987;

FEATURES             Location/Qualifiers
     source          1..2092
                     /organism="unknown"
BASE COUNT      473 a    605 c    593 g    421 t
ORIGIN

  Query Match             10.7%;  Score 135;  DB 55;  Length 2092;
  Best Local Similarity   77.8%;  Pred. No. 1.36e-103;
  Matches  189;  Conservative    0;  Mismatches   54;  Indels    0;  Gaps    0;

Db       832 ggagactcgctactgtgcagtgtgcaatgactatgcttcaggctaccattatggagtctg 891
             |||  ||| |  ||| || || ||||  || || || || || || || ||||||||||||
Qy       273 GGATGCTCACTTCTGCCCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTG 332

Db       892 gtcctgtgagggctgcaaggccttcttcaagagaagtattcaaggacataaagactatat 951
             |||  ||||| || || |||||||| || || |||||  |||||||||||||| || |||||
Qy       333 GTCGTGTGAAGGATGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATAT 392

Db       952 gtgtccagccaccaaccagtgccaccattgataaaaacaggagaagagctgccaggcctg 1011
              |||||||| || || ||||| || || ||||||||| || | ||||||||||||||||||
Qy       393 TTGTCCAGCTACAAATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTG 452

Db      1012 ccggctcggcaaatgctacgaagtgggaatgatgaaaggtgggatccgaaaagaccgaag 1071
             ||| || || || || ||||||||||||||| ||||  ||||     || | |||  ||| |
Qy       453 CCGACTTCGGAAGTGTTACGAAGTGGGAATGGTGAAGTGTGGCTCCGGAGAGAGAGATG 512

Db      1072 agg 1074
              ||
Qy       513 TGG 515


Search completed: Thu Dec 18 19:46:01 1997
Job time : 1077 secs.
```

```
*****************************************************************

    MPSRCH                                              (TM)

*****************************************************************

         Release 2.1D John F  Collins, Biocomputing Research Unit.
         Copyright (c) 1993, 1994, 1995  University of Edinburgh, U K
                 Distribution rights by IntelliGenetics  Inc.

MPsrch_nn   n a  -  n a  database search  using Smith-Waterman algorithm

Run on:            Thu Dec 18 17:18:36 1997;  MasPar time 1194 04 Seconds
                                               1318 993 Million cell updates/sec
Tabular output not generated

Title:             >US-08-826-361A-1
Description:       (1-1434) from US08826361A seq
Perfect Score:     1434
N.A. Sequence.         1 ATGAATTACAGCATTCCTAG          AGAACCCACAGTCTCAGTGA 1434
     Comp                TACTTAATGTCGTAAGGATC          TCTTGGGTGTCAGAGTCACT

Scoring table:     TABLE default
                   Gap 6

Nmatch   STD :     Dbase 0;  Query 0

Searched:          362067 seqs, 549128275 bases x 2

Post-processing    Minimum Match 0%
                   Listing first 45 summaries

Database:          embl-new3
                     1:BCT 2:FUN 3:GEN1 4:GEN2 5:HTG1 6:HTG2 7:HUM 8:INV
                     9:ORG 10:MAM 11:VRT 12:PLN 13:PRO 14:ROD 15:SYN 16:UNC
                     17:VIR
Database:          genbank99
                     18:BCT1 19:BCT2 20:BCT3 21:BCT4 22:BCT5 23:BCT6 24:BCT7
                     25:BCT8 26:BCT9 27:BCT10 28:BCT11 29:GEN1 30:GEN2
                     31:GEN3 32:HTG1 33:HTG2 34:HTG3 35:INV1 36:INV2 37:INV3
                     38:INV4 39:INV5 40:INV6 41:INV7 42:INV8 43:INV9 44:INV10
                     45:INV11 46:MAM1 47:MAM2 48:MAM3 49:VRT1 50:VRT2 51:VRT3
                     52:VRT4 53.PAT1 54.PAT2 55.PAT3 56.PAT4 57.PAT5 58:PHG
                     59:PLN1 60:PLN2 61:PLN3 62:PLN4 63:PLN5 64:PLN6 65:PLN7
                     66:PLN8 67:PLN9 68:PLN10 69:PLN11 70:PRI1 71:PRI2
                     72:PRI3 73:PRI4 74:PRI5 75:PRI6 76:PRI7 77:PRI8 78:PRI9
                     79:PRI10 80:PRI11 81:PRI12 82:PRI13 83:PRI14 84:PRI15
                     85:ROD1 86:ROD2 87:ROD3 88:ROD4 89:ROD5 90:ROD6 91:ROD7
                     92:ROD8 93:STS 94:SYN 95:UNA 96:VRL1 97.VRL2 98.VRL3
                     99:VRL4 100:VRL5 101:VRL6 102:VRL7 103:VRL8 104.VRL9
                     105:VRL10
Database:          genbank-new3
                     106:BCT 107:GEN1 108:GEN2 109:HTG1 110:HTG2 111:INV
                     112:MAM 113:VRT 114:PHG 115:PLN 116:PRI1 117:PRI2
                     118:ROD 119:SYN 120:UNA 121:VRL
Database:          u-embl50_99
                     122:part1

Statistics:        Mean 11.436;  Variance 4.445;  scale 2.572

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES
```

```
Result         Query
  No.  Score   Match Length DB  ID           Description            Pred. No.
-------------------------------------------------------------------------------
    1   1434   100.0   1560 76  HSRNAERB     H sapiens mRNA for es  0.00e+00
    2   1412    98.5   2255 42  RNU57439     Rattus norvegicus est  [illegible]
    3    261    18.2   [illegible] 73  CJERB C.japonicus mRNA for es  [illegible]
    4    217    15.1   1729 51  [illegible]  Killifish mRNA for es  2.58e-199
    5    203    14.2   1788 4?  SSESTREC     S.scrofa mRNA for est  1.71e-174
    6    201    14.0   1966 47  OAESTREC     O.aries mRNA for estr  2.21e-172
    7    201    14.0   2085 50  GGERR        Chicken mRNA for est:  2.21e-172
    8    197    13.7   2092 55  I08538       Sequence 1 from Paten  3.68e-168
    9    197    13.7   2090 90  HUMERMCI     Human estrogen recept  3.68e-168
   10    197    13.7   4963 56  I15368       Sequence 16 from pate  3.68e-168
   11    197    13.7   6450 74  HSERR        Human mRNA for oestro  3.68e-168
   12    189    13.2   2792 52  ZFIER        Poephila guttata estr  9.84e-160
   13    184    12.8   1880 91  RNESTROR     Rat mRNA for estrogen  1.77e-154
   14    184    12.8   2090 91  RNERR        Rat mRNA for estrogen  1.77e-154
   15    184    12.8   2090 92  S52128       estrogen receptor [ra  1.77e-154
   16    178    12.4   2078 88  MUSESTRU     Mouse estrogen recept  3.50e-148
   17    175    12.2    586 79  HSZ75126     H sapiens mRNA for es  4.96e-145
   18    172    12.0   1662 52  SSOESREC     S.salar mRNA for oest  6.68e-142
   19    170    11.9   1279 52  SMOERA       Rainbow trout estroge  9.22e-140
   20    141     9.8    923 51  S79925       estrogen receptor-rep  1.03e-109
   21    140     9.8   2106 78  HSU47678     Human 80 kDa estrogen  1.10e-108
   22    136     9.5   2449 52  XELESTRREC   Xenopus laevis estrog  1.43e-104
   23    136     9.5   2479 52  XELESTRO     Xenopus laevis estrog  1.43e-104
   24    136     9.5   2533 52  XELER3       Xenopus laevis estrog  1.43e-104
   25    135     9.4    477 46  BTU64962     Bos taurus estrogen r  1.52e-103
   26    133     9.3   2619 80  HUMERVBCA    Human estrogen recept  1.70e-101
   27    109     7.6    735 47  OAU30299     Ovis aries estrogen r  3.96e-77
   28     78     5.4    323 74  HSEST        H. sapiens mRNA for e  1.32e-46
   29     78     5.4    323 117  HSEST        H. sapiens mRNA for e  1.32e-46
   30     77     5.4    400 46  BTERREC      B taurus mRNA for est  1.30e-45
   31     74     5.2    365 51  OAESREXN4    O.aureus estrogen rec  8.78e-43
   32     71     5.0   3400 57  I25545       Sequence 1 from paten  6.15e-40
   33     68     4.7    363 84  S71040       estradiol receptor [M  4.07e-37
   34     68     4.7   2153 55  I08706       Sequence 13 from Pate  4.07e-37
   35     68     4.7   2153 76  HSSTHOR2     Human mRNA for steroi  4.07e-37
   36     66     4.6   1344 92  S92459       mERR 2-estrogen recpt  3.01e-35
   37     66     4.6   1344 122  MMERR2GEN    M.musculus mRNA for e  3.01e-35
   38     64     4.5   2265 85  MMGRIA       Mouse mRNA for gluco  [illegible]
   39     63     4.4   2331 49  RABPGR       Rabbit progesterone r  1.82e-32
   40     63     4.4   6244 56  I15367       Sequence 15 from pate  1.82e-32
   41     62     4.3   2376 10  ANU87952     Aotus nancymaae gluco  1.52e-31
   42     62     4.3   2376 10  SOU87953     Saguinus oedipus gluc  1.52e-31
   43     62     4.3   2376 119  SOU87953     Saguinus oedipus gluc  1.52e-31
   44     62     4.3   2376 119  ANU87952     Aotus nancymaae gluc  1.52e-31
   45     62     4.3   2575 85  MMGCRR       Mouse mRNA for glucoc  1.52e-31


                               ALIGNMENTS

RESULT      1
LOCUS       HSRNAERB      1560 bp    RNA               PRI       05-FEB-1996
DEFINITION  H.sapiens mRNA for estrogen receptor.
ACCESSION   X99101
NID         g1518262
KEYWORDS    estrogen receptor.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata, Eutheria, Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1560)
  AUTHORS   Mosselman,S , Polman,J. and Dijkema,R.
  TITLE     ER beta: identification and characterization of a novel human
            estrogen receptor
  JOURNAL   FEBS Lett. 392 (1), 49-53 (1996)
  MEDLINE   96354875
REFERENCE   2  (bases 1 to 1560)
  AUTHORS   Mosselman,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-JUL-1996) S  Mosselman, N V Organon, Biotechnology
            and Biochemistry, PO box 20 oss, Molenstraat 110 oss, 5340 BH,
                     NETHERLANDS
FEATURES             Location/Qualifiers
     source          1..1560
                     /organism="Homo sapiens"
                     /tissue_type="testis"
                     /dev_stage="adult"
                     /clone_lib="Clontech, HL1010B"
                     /clone_lib="Clontech; 7414-1"
     CDS             19..1452
                     /codon_start=1
                     /product="estrogen receptor beta"
                     /db_xref="PID:e255390"
                     /db_xref="PID:g1518263"
                     /translation="MNYSIPSNVTNLEGGPGRQTTSPNVLWPTPGHLSPLVVHRQLSH
                     LYAEPQKSPWCEARSLEHTLPVNRETLKRKVSGNRCASPVTGPGSKRDAHFCAVCSDY
                     ASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRKCYEVG
                     MVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQLVLT
                     LLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFDQVR
                     LLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSRF
                     RELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKSG
                     ISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNAHVLRGC
                     KSSITGSECSPAEDSKSKEGSQNPQSQ"
BASE COUNT      367 a    408 c    445 g    340 t
ORIGIN

  Query Match          100.0%;  Score 1434;  DB 76;  Length 1560;
  Best Local Similarity 100.0%;  Pred. No. 0.00e+00;
  Matches 1434;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Db         19 atgaattacagcattcccagcaatgtcactaacttggaaggtgggcctggtcggcagacc 78
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy          1 ATGAATTACAGCATTCCCAGCAATGTCACTAACTTGGAAGGTGGGCCTGGTCGGCAGACC 60

Db         79 acaagcccaaatgtgttgtggccaacacctgggcacctttctcctttagtggtccatcgc 138
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy         61 ACAAGCCCAAATGTGTTGTGGCCAACACCTGGGCACCTTTCTCCTTTAGTGGTCCATCGC 120

Db        139 cagttatcacatctgtatgcggaacctcaaaagagtccctggtgtgaagcaagatcgcta 198
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        121 CAGTTATCACATCTGTATGCGGAACCTCAAAAGAGTCCCTGGTGTGAAGCAAGATCGCTA 180

Db        199 gaacacaccttacctgtaaacagagagacactgaaaaggaaggttagtgggaaccgttgc 258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        181 GAACACACCTTACCTGTAAACAGAGAGACACTGAAAAGGAAGGTTAGTGGGAACCGTTGC 240

Db        259 gccagccctgttactggtccaggttcaaagagggatgctcacttctgcgctgtctgcagc 318
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        241 GCCAGCCCTGTTACTGGTCCAGGTTCAAAGAGGGATGCTCACTTCTGCGCTGTCTGCAGC 300

Db        319 gattacgcatcgggatatcactatggagtctggtcgtgtgaaggatgtaaggcctttttt 378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        301 GATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATGTAAGGCCTTTTTT 360

Db        379 aaaagaagcattcaaggacataatgattatatttgtccagctacaaatcagtgtacaatc 438
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        361 AAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCAGTGTACAATC 420

Db        439 gataaaaaccggcgcaagagctgccaggcctgccgacttcggaagtgttacgaagtggga 498
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        421 GATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTACGAAGTGGGA 480

Db        499 atggtgaagtgtggctcccggagagagagatgtgggtaccgccttgtgcggagacagaga 558
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        481 ATGGTGAAGTGTGGCTCCCGGAGAGAGAGATGTGGGTACCGCCTTGTGCGGAGACAGAGA 540

Db        559 agtgccgacgagcagctgcactgtgccggcaaggccaagagaagtggcggccacgcgccc 618
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        541 AGTGCCGACGAGCAGCTGCACTGTGCCGGCAAGGCCAAGAGAAGTGGCGGCCACGCGCCC 600

Db        619 cgagtgcgggagctgctgctggacgccctgagccccgagcagctagtgctcaccctcctg 678
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        601 CGAGTGCGGGAGCTGCTGCTGGACGCCCTGAGCCCCGAGCAGCTAGTGCTCACCCTCCTG 660

Db        679 gaggctgagccgccccatgtgctgatcagccgccccagtgcgcccttcaccgaggcctcc 738
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        661 GAGGCTGAGCCGCCCCATGTGCTGATCAGCCGCCCCAGTGCGCCCTTCACCGAGGCCTCC 720

Db        739 atgatgatgtccctgaccaagttggccgacaaggagttggtacacatgatcagctgggcc 798
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        721 ATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTACACATGATCAGCTGGGCC 780

Db        799 aagaagattcccggctttgtggagctcagcctgttcgaccaagtgcggctcttggagagc 858
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        781 AAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAAGTGCGGCTCTTGGAGAGC 840

Db        859 tgttggatggaggtgttaatgatggggctgatgtggcgctcaattgaccaccccggcaag 918
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        841 TGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCAATTGACCACCCCGGCAAG 900

Db        919 ctcatctttgctccagatcttgttctggacagggatgaggggaaatgcgtagaaggaatt 978
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        901 CTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGGAAATGCGTAGAAGGAATT 960

Db        979 ctggaaatctttgacatgctcctggcaactacttcaaggtttcgagagttaaaactccaa 1038
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        961 CTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTTCGAGAGTTAAAACTCCAA 1020

Db       1039 cacaaagaatatctctgtgtcaaggccatgatcctgctcaattccagtatgtaccctctg 1098
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1021 CACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAATTCCAGTATGTACCCTCTG 1080

Db       1099 gtcacagcgacccaggatgctgacagcagccggaagctggctcacttgctgaacgccgtg 1158
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1081 GTCACAGCGACCCAGGATGCTGACAGCAGCCGGAAGCTGGCTCACTTGCTGAACGCCGTG 1140

Db       1159 accgatgctttggtttgggtgattgccaagagcggcatctcctcccagcagcaatccatg 1218
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1141 ACCGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATCTCCTCCCAGCAGCAATCCATG 1200

Db       1219 cgcctggctaacctcctgatgctcctgtcccacgtcaggcatgcgagtaacaagggcatg 1278
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1201 CGCCTGGCTAACCTCCTGATGCTCCTGTCCCACGTCAGGCATGCGAGTAACAAGGGCATG 1260

Db       1279 gaacatctgctcaacatgaagtgcaaaaatgtggtcccagtgtatgacctgctgctggag 1338
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1261 GAACATCTGCTCAACATGAAGTGCAAAAATGTGGTCCCAGTGTATGACCTGCTGCTGGAG 1320

Db       1339 atgctgaatgcccacgtgcttcgcgggtgcaagtcctccatcacggggtccgagtgcagc 1398
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1321 ATGCTGAATGCCCACGTGCTTCGCGGGTGCAAGTCCTCCATCACGGGGTCCGAGTGCAGC 1380

Db       1399 ccggcagaggacagtaaaagcaaagagggctcccagaacccacagtctcagtga 1452
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1381 CCGGCAGAGGACAGTAAAAGCAAAGAGGGCTCCCAGAACCCACAGTCTCAGTGA 1434


RESULT      2
LOCUS       RNU57439     2555 bp    mRNA             ROD       12-JUN-1996
DEFINITION  Rattus norvegicus estrogen receptor beta mRNA, complete cds.
ACCESSION   U57439
NID         g1373280
KEYWORDS    .
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
            Murinae; Rattus.
REFERENCE   1  (bases 8 to 2555)
  AUTHORS   Kuiper,G.G., Enmark,E., Pelto-Huikko,M., Nilsson,S. and
            Gustafsson,J.A.
  TITLE     Cloning of a novel receptor expressed in rat prostate and ovary
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 93 (12), 5925-5930 (1996)
  MEDLINE   96234066
REFERENCE   2  (bases 1 to 2555)
  AUTHORS   Enmark,E.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-MAY-1996) Eva I  Enmark  Center for Biotechnology
            NOVUM, Karolinska, Institute, NOVUM, Huddinge, S-141 57, Sweden
FEATURES             Location/Qualifiers
     source          1..2555
                     /organism="Rattus norvegicus"
                     /strain="Sprague Dawley"
                     /sex="male"
                     /tissue_type="prostate"
     CDS             417..1874
                     /note="nuclear receptor superfamily member"
                     /codon_start=1
                     /product="estrogen receptor beta"
                     /db_xref="PID:g1373281"
                     /translation="MTFYSPAVMNYSVPGSTSNLDGGPVRLSTSPNVLWPTSGHLSPL
                     ATHCQSSLLYAEPQKSPWCEARSLEHTLPVNRETLKRKLSGSSCASPVTSPNAKRDAH
                     FCPVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACR
                     LRKCYEVGMVKCGSRRERCGYRIVRRQRSSSEQVHCLSKAKRNGGHAPRVKELLLSTL
                     SPEQLVLTLLEAEPPNVLVSRPSMPFTEASMMMSLTKLADKELVHMIGWAKKIPGFVE
                     LSLLDQVRLLESCWMEVLMVGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDM
                     LLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLASANQEAESSRKLTHLLNAVTDAL
                     VWVIAKSGISSQQQSVRLANLLMLLSHVRHISNKGMEHLLSMKCKNVVPVYDLLLEML
                     NAHTLRGYKSSISGSECSSTEDSKNKESSQNLQSQ"
     misc_feature    726..923
                     /note="encodes DNA binding domain"
BASE COUNT      630 a    652 c    660 g    613 t
ORIGIN

  Query Match            70.6%;  Score 1012;  DB 92;  Length 2555;
  Best Local Similarity  85.3%;  Pred. No. 0.00e+00;
  Matches 1223;  Conservative    0;  Mismatches 211;  Indels    0;  Gaps    0;

Db     441 atgaactacagtgtttccggcagcaacagtaacctggacggtggcctgtccgactgagc 500
           ||||| |||||  ||||| |||     || |||| |||| |||||||||   || | || |
Qy       1 ATGAATTACAGCATTCCCAGCAATGTCACTAACTTGGAAGGTGGGCCTGGTCGGCAGACC 60

Db     501 accagcccaaatgtgttgtggccaacacctgggcacctgtctctttagcgacccattgc 560
           ||||||||||||||| | ||||||||  |||||||||| ||||||||||| |  |||| ||
Qy      61 ACAAGCCCAAATGTGTTGTGGCCAACACCTGGGCACCTTTCTCCTTTAGTGCTCCATCGC 120

Db     561 caatcatcgctcctctatgcagaacctcaaaagagtccttggtgtgaagcaagatcacta 620
           || | ||| |  || ||||| ||||||||||||||||| |||||||||||||||| |||
Qy     121 CAGTTATCACATCTGTATGCGGAACCTCAAAAGAGTCCCTGGTGTGAAGCAAGATCGCTA 180

Db     621 gagcacaccttacctgtaaacagagagacactgaagaggaagcttagtgggagcagttgt 680
           || ||||||||||||||||||||||||||||||||| |||||| ||||||||| | ||||
Qy     181 GAACACACCTTACCTGTAAACAGAGAGACACTGAAAAGGAAGCTTAGTGGGAACCCTTGC 240

Db     681 gccagccctgttactagttcaaacgcaaagagggatgctcacttctgcgccgtctgcagc 740
           ||||||||||||||| |||||     ||||||||||||||||||||||| | |||||||||
Qy     241 GCCAGCCCTGTTACTGGTCCAGGTTCAAAGAAGGATGCTCACTTCTGCGCTGTCTGCAGC 300

Db     741 gattatgcatctgggtatcattacggcgtttggtcatgtgaaggatgtaaggcctttttt 800
           ||||| ||||| || ||||| || || || ||||| ||||||||||||||||||||||||
Qy     301 GATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATGTAAGGCCTTTTTT 360

Db     801 aaaagaagcattcaaggacataatgattatatctgtccagccacgaatcagtgtaccata 860
           ||||||||||||||||||||||||||||||||| |||||||| || ||||||||||| ||
Qy     361 AAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCAGTGTACAATC 420

Db     861 gacaagaaccggcgtaaaagctgccaggcctgccgacttcgaagtgttatgaagtagga 920
           || || |||||||| || |||||||||||||||||||||||| |||||||| ||||| |||
Qy     421 GATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTACCAAGTGGGA 480

Db     921 atggtcaagtgtggatccaggagagaaacggtgtgggtaccgtatagtgggaggcagaga 980
           ||||| |||||||| ||| |||||||    | |||||||||||    | |||||||| ||||||
Qy     481 ATGGTGAAGTGTGGCTCCCGGAGAGAGAGATGTGGGTACCGCCTTGTGCCCAGACAGAGA 540

Db     981 agttctagcgagcaggtacactgcctgagcaaagccaagagaaacggtgggcatgcaccc 1040
           |||  |    ||||||| | |||||       |||| |||||||||  || || || || |||
Qy     541 AGTGCCGACGAGCAGCTGCACTGTGCCGGCAAGGCCAAGAGAAGTGGCGGCCACGCGCCC 600

Db    1041 cgggtgaaggagctactgctgagcaccttgagtccagagcaactggtgctcaccctcctg 1100
           || |||   |||||| ||||||    | || |||| || ||||| || ||||||||||||||
```

```
Qy     601 CGAGTGCGGGAGCTGCTGCTGGACGCCCTGAGCCCCGAGCAGCTAGTGCTCACCCTCCTG 660

Db    1101 gaagctgaaccacctaatgtgctggtgagccgtcccagcatgcccttcaccgaagcctcc 1160
           || ||||| || ||| ||||||| | ||||| |||||    ||||||||||||| |||||
Qy     661 GAGGCTGAGCCGCCCCATGTGCTGATCAGCCGCCCCAGTGCGCCCTTCACCGAGGCCTCC 720

Db    1161 atgatgatgtccctcactaagctggcggacaaggaactggtgcacatgattggctgggcc 1220
           ||||||||||||||| || ||| |||| ||||||||   |||| |||||||   ||| ||||
Qy     721 ATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTACACATGATCAGCTGGGCC 780

Db    1221 aagaaaatccctggctttgtggagctcagcctgttggaccaagtccgcctcttagaaagc 1280
           ||| ||| ||| || ||||||||||||||||||||||||| |||||||| |||| |   |
Qy     781 AAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAAGTGCGGCTCTTGGAGAGC 840

Db    1281 tgctggatggaggtgctaatggtgggactgatgtggcgctccatcgaccaccccggcaag 1340
           || |||||||||||| |||||| |||| |||||||||||||| || ||||||||||| ||
Qy     841 TGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCAATTGACCACCCCGGCAAG 900

Db    1341 ctcatttttgctcccgacctcgttctggacagggatgagggaaagtgcgtagaaggaatt 1400
           ||||| || ||||| || || ||||||||||||||||||||||||| ||||| ||||||| ||
Qy     901 CTCATCTTGGCTCCAGATCTTGTTCTGGACAGGGATGAGGGAAAATGCGTAGAAGGAATT 960

Db    1401 ctggaaatctttgacatgctcctggcgacgacgtcaaggttccgtgagttaaaactccag 1460
           ||||||||||||||||||||||||||| || || |||||||| || ||| |||||||||||
Qy     961 CTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTTCGAGAGTTAAAACTCCAA 1020

Db    1461 cacaaggagtatctctgtgtgaaggccatgatcctcctcaactccagtatgtaccctttg 1520
           ||||| || ||||||||||| |||||||||||||||| ||||| ||||||||||||||   ||
Qy    1021 CACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAATTCCAGTATGTACCCTCTG 1080

Db    1521 gcttctgcaaaccaggaggcagaaagtagccggaagctgacacacctactaaacgcggtg 1580
           |    | || | |||||| || || || ||||||||||||| | || | ||||||| |||
Qy    1081 GTCACAGCGACCCAGGATGCTGACAGCAGCCGGAAGCTGGCTCACTTGCTGAACGCCGTG 1140

Db    1581 acagatgccctggtctgggtgattgcgaagagtggtatctcctcccagcagcagtcagtc 1640
           || ||||||   ||||  |||||||||||| ||||| || ||||||||||||||||  ||  |
Qy    1141 ACCGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATCTCCTCCCAGCAGCAATCCATC 1200

Db    1641 cgactggccaacctcctgatgcttctttctcacgtcaggcacatcagtaacaagggcatg 1700
           || ||||| ||||||||||||||| || || ||||||||||     ||||||||||||||
Qy    1201 CGCCTGGCTAACCTCCTGATGCTCCTGTCCCACGTCAGGCATGCGAGTAACAAGGGCATG 1260

Db    1701 gaacatctgctcagcatgaagtgcaaaaatgtggtcccggtgtatgacctgctgctggag 1760
           ||||||||||||| ||||||||||||||||||||||||| ||||||||||||||||||||
Qy    1261 GAACATCTGCTCAACATGAAGTGCAAAAATGTGGTCCCAGTGTATGACCTGCTGCTGGAG 1320

Db    1761 atgctgaatgctcacacgcttcgagggtacaagtcctcaatctcggggtctgagtgcagc 1820
           ||||||||||| |||   |||||| |||| ||||||||| ||| ||||||| ||||||||
Qy    1321 ATGCTGAATGCCCACGTGCTTCGCGGGTGCAAGTCCTCCATCACGGGGTCCGAGTGCAGC 1380

Db    1821 tcaacagaggacagtaagaacaaagagagctcccagaacctacagtctcagtga 1874
             |  ||||||||||||| | ||||||| |||||||||||| ||||||||||||||
Qy    1381 CCGGCAGAGGACAGTAAAAGCAAAGAGGGCTCCCAGAACCCACAGTCTCAGTGA 1434


RESULT     3
LOCUS       CJERB         332 bp    RNA             PRI       14-NOV-1996
DEFINITION  C.jacchus mRNA for estrogen receptor beta.
ACCESSION   Y09372
NID         g1563661
KEYWORDS    DNA binding; estrogen receptor; steroid hormone receptor;
            transactivation.
SOURCE      common marmoset.
  ORGANISM  Callithrix jacchus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Platyrrhini; Callitrichidae;
            Callithrix.
REFERENCE   1  (bases 1 to 332)
  AUTHORS   Gaughan,J.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 332)
  AUTHORS   Gaughan,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-NOV-1996) J. Gaughan, Medical Research Council,
            Centre for Reproductive Biology, 37 Chalmers Street, Edinburgh, EH3
            9EW, UK
COMMENT     Related entries: X99101 & U57439
            Related entries: X99101 & U57439.
FEATURES             Location/Qualifiers
     source          1..332
                     /organism="Callithrix jacchus"
                     /tissue_type="testis"
                     /tissue_type="ovary"
                     /tissue_type="prostate"
                     /clone="pGEM-T vector"
     CDS             18..>332
                     /function="steroid hormone receptor"
                     /codon_start=1
                     /product="estrogen receptor beta"
                     /db_xref="PID:e291463"
                     /db_xref="PID:g1669662"
                     /translation="MNYSIPSSVTNLEAGPGRQTTSPNMLWSTPGHLSPLAVHHQLSH
                     LYAEPQKTPWCEARSLEHTLPVSRETLKRKVSGNHCASPVIGPSSKRDAHFCAVCSDY
                     AAE"
BASE COUNT       91 a     84 c     77 g     80 t
ORIGIN

  Query Match             18.2%;  Score 261;  DB 70;  Length 332;
  Best Local Similarity   92.2%;  Pred. No. 3.39e-236;
  Matches  285;  Conservative    0;  Mismatches   24;  Indels    0;  Gaps    0;

Db       18 atgaattacagcattcccagcagtgttactaacttggaagcggggcctggtcggcagatc 77
            |||||||||||||||||||||| ||| |||||||||||||| |||||||||||||||| |
Qy        1 ATGAATTACAGCATTCCCAGCAATGTCACTAACTTGGAAGCCGGGCCTGGTCGGCAGACC 60

Db       78 acaagcccaaatatgctgtggtcaacacctgggcacctttctcctttagctgtccatcac 137
            |||||||||||| || ||||| |||||||||||||||||||||||||||  ||||||| |
Qy       61 ACAAGCCCAAATGTGTTGTGGCCAACACCTGGGCACCTTTCTCCTTTAGTGGTCCATCGC 120

Db      138 cagttatcacatctgtatgcagaacctcaaaagactccctggtgtgaagcaagatcacta 197
            |||||||||||||||||||| |||||||||||| |||||||||||||||||||||| |||
Qy      121 CAGTTATCACATCTGTATGCGGAACCTCAAAAGAGTCCCTGGTGTGAAGCAAGATCGCTA 180

Db      198 gaacatacccttacctgtaagcagagagacactgaaaaggaaagttagtgggaaccattgt 257
            ||||| ||||||||||||| ||||||||||||||||||||| ||||||||||||| |||
Qy      181 GAACACACCTTACCTGTAAACAGAGAGACACTGAAAAGGAAGGTTAGTGGGAACCGTTGC 240

Db      258 gccagccctgttactggtccaggttcaaagagggatgctcacttctgtgcagtctgcagc 317
            |||||||||||||||||||| ||||||||||||||||||||||||||| || ||||||||
Qy      241 GCCAGCCCTGTTACTGGTCCAGGTTCAAAGAGGGATGCTCACTTCTGCCCAGTCTGCAGC 300

Db      318 gattatgca 326
            ||||| |||
Qy      301 GATTACGCA 309


RESULT      4
LOCUS       ORZMER        1728 bp    mRNA             VRT       22-DEC-1994
DEFINITION  Killifish mRNA for estrogen receptor, complete cds.
ACCESSION   D28954
NID         g473547
KEYWORDS    estrogen receptor.
SOURCE      Oryzias sp. (strain d-rR) (library: lambda ZAP II) liver cDNA to
            mRNA.
  ORGANISM  Oryzias sp.
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Actinopterygii; Neopterygii; Teleostei; Euteleostei;
            Acanthopterygii; Atherinomorpha; Cyprinodontiformes;
            Adrianichthyoidei; Oryziidae; Oryzias.
REFERENCE   1  (bases 1 to 1728)
  AUTHORS   Okada,H.
  JOURNAL   Unpublished (1994)
COMMENT     Submitted (15-Mar-1994) to DDBJ by:
            Hisashi Okada
            Center for Gene Science
            Hiroshima University
            1-4-2 Kagamiyama, Higashi-Hiroshima
            Hiroshima 724
            Japan
            Phone: 0824-24-6283
            Fax:   0824-22-7184.
FEATURES             Location/Qualifiers
     source          1..1728
                     /organism="Oryzias sp."
                     /strain="d-rR"
                     /tissue_type="liver"
                     /clone_lib="lambda ZAP II"
     CDS             1..1728
                     /gene="MER"
                     /codon_start=1
                     /product="estrogen receptor"
                     /db_xref="PID:g602326"
                     /translation="MYPEESRGSGGVLLWTTWKGRTTMPAPNFATTHLYSQSSTGYYS
                     APLEINGPPSEGSLQSLGSGPTSPLVFVPSSPRLSPFMHPPSHHYLETTSIPVYRSSH
                     QGASREDQCGSREDTCSLGELGAGAGAGGFEMAKDTRFCAVCSDYASGYHYGVWSCEG
                     CKAFFKRSIQGHNDYMCPATNQCTIDRNRRKSCQACRLRKCYEVGMMKGGVRKDRIRI
                     LRRDKRRTGVGDGDKVVKGQEHKTVHYDGRKRSSTGGGGGGGGRLSVTSIPPEQVLL
                     LLQGAEPPILCSRQKLSRPYTEVTMMTLLTSMADKELVHMIAWAKKLPGFLQLSLHDQ
                     VLLLESSWLEVLMIGLIWRSIHCPGKLIFAQDLILDRNEGDCVEGMTEIFDMLLATAS
                     RFRVLKLKPEEFVCLKAIILLNSGAFSFCTGTMEPLHNSAAVQSMLDTITDALIHYIS
                     QSGYLAQEQARRQAQLLLLLSHIRHMSNKGMEHLYSMKCKNKVPLYDLLLEMLDAHRL
                     HHPVRAPQSLSQVDRDPPSTSSGGGTAPGSTSASRGRILSPSRGPPAPSVLQYGGSR
                     PDCIPALQD"
BASE COUNT      379 a    513 c    496 g    340 t
ORIGIN

  Query Match             15.1%;  Score 217;  DB 51;  Length 1728;
  Best Local Similarity   68.2%;  Pred. No. 2.58e-189;
  Matches  434;  Conservative    0;  Mismatches  199;  Indels    3;  Gaps    1;

Db      883 ccgtacaccgaggtcaccatgatgaccctgctcaccagcatggcagacaaggagctggtc 942
            || | |||||||| | |||||||||      || ||||    |||| ||||||||| ||||
Qy      703 CCCTTCACCGAGGCCTCCATGATGATCTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTA 762

Db      943 cacatgatcgcctgggccaagaagctcccaggtttcctgagctgtccctgcaggatcaa 1002
            |||||||||  ||||||||||||| | || || ||| || ||||    ||||   ||| ||
Qy      763 CACATGATCAGCTGGGCCAAGAAGATCCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAA 822

Db     1003 gtgctgctgctggagagctcgtggctggaggtgctcatgatcggcctcatttggaggtcc 1062
            |||| |||  ||||||||  ||| |||||||| | ||||| || || || ||| | ||
Qy      823 GTGCGGCTCTTGGAGAGCTGCTTGGATGGACGTGTTAATGATGGGCTGATGTGGCGCTCA 882

Db     1063 atccactgtcccgggaagctcatctttgcacaagacctcatcctggacaggaatgaggga 1122
            ||  ||   ||||| ||||||||||||| | ||| ||   | ||||||||| ||||||
Qy      883 ATTGACCACCCCGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGG 942

Db     1123 gactgcgtggaaggcatgacggagatcttcgacatgctgctggccactgcttcccgcttc 1182
             | ||||| ||||| ||    ||| ||||| |||||||| ||||| ||| ||||   | ||
Qy      943 AAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTT 1002

Db     1183 cgtatgctcaaactcaaacctgaggagttcgtgtgcctcaaagctatcattttactcaac 1242
            || | | | |||||| |||    | ||||    ||||   |||| || || ||   | |||||
Qy     1003 CGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAAT 1062

Db     1243 tccggtgctttttctttctgcaccggcaccatggagccacttcacaacagcgcggcggtt 1302
            ||| ||    |    || |    ||| |   |||   |||    |    | ||| ||||    |   | |
Qy     1063 TCCAGTATGTACCCTCTGGTCACAGCGACCCAGGATGC---TGACAGCAGCCGGAAGCTG 1119

Db     1303 cagagcatgctggacaccatcacagacgcactcatccactacatcagtcagtcaggttac 1362
               | ||||| || || | || || ||  |  ||         ||      ||    ||     |
Qy     1120 GCTCACTTGCTGAACGCCGTGACCGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATC 1179

Db     1363 ttggcccaggagcaggcgaggcggcaggcgcagctgctgctgctgctgtcccacatcagg 1422
            |    ||||| ||||    | |    || | |||   | || ||   |||| || |||||| |||||
Qy     1180 TCCTCCCAGCAGCAATCCATGCGCCTGGCTAACCTCCTGATGCTCCTGTCCCACGTCAGG 1239

Db     1423 cacatgagcaacaaaggcatggagcacctctacagcatgaagtgcaagaacaaggtccct 1482
            ||     ||| ||||| |||||||| || ||     || ||||||||||| ||      |||||
Qy     1240 CATGGACTAACAAGGGCATGAACATCTGCTCAACATGAAGTGCAAAAATGTGCCTCCCA 1299
```

```
Db   1483 ctttatgacctcctactggagatgctcaatgcccac 1518
               | ||||||| || | |||||||||  ||||||||
Qy   1300 GTGTATGACCTGCTGCTGGAGATGCTGAATGCCCAC 1335


RESULT     5
LOCUS       SSESTREC     1788 bp    RNA             MAM       15-SEP-1994
DEFINITION  S.scrofa mRNA for estradiol receptor.
ACCESSION   Z37167
NID         g587554
KEYWORDS    estradiol receptor.
SOURCE      pig.
  ORGANISM  Sus scrofa
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Artiodactyla; Suiformes; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 1788)
  AUTHORS   Bokenkamp,D., Jungblut,P.W. and Thole,H.H.
  TITLE     The C-terminal half of the porcine estradiol receptor contains no
            post-translational modification: determination of the primary
            structure
  JOURNAL   Mol. Cell. Endocrinol. 104 (2), 163-172 (1994)
  MEDLINE   95080454
REFERENCE   2  (bases 1 to 1788)
  AUTHORS   Boekenkamp,D.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-SEP-1994) Dirk Boekenkamp, Max-Planck-Institut fuer
            experimentelle Endokrinologie, Feodor-Lynen Strasse 7, Hannover,
            30625, Germany
FEATURES             Location/Qualifiers
     source          1..1788
                     /organism="Sus scrofa"
                     /tissue_type="Uterus"
                     /sex="Female"
     CDS             1..1788
                     /codon_start=1
                     /product="estradiol receptor"
                     /db_xref="PID:g587555"
                     /translation="MTMTLHTKASGMALLHQIQANELEPLNRPQLKIPLERPLGEVYV
                     DSSKPAVYNYPEGAAYDFNAAAAASAPVYGQSGLAYGPGSEAAAFGANGLGGFQPLNS
                     VSPSPLVLLHPPPQLSPFLHPHGQQVPYYLENEPSGYAVREAGPPAFYRPNSDNRRQG
                     GRERLASTSDKGSMAMESAKETRYCAVCNDYASGYHYGVWSCEGCKAFFKRSIQGHND
                     YMCPATNQCTIDKNRRKSCQACRLRKCYEVGMMKGGIRKDRRGGRMLKHKRQRDDGEG
                     RNEAVPPGDMRSANLWPSPLLIKHTKKNSPVLSLTADQMISALLEAEPPIIYSEYDPT
                     RPLSEASMMGLLTNLADRELVHMINWAKRVPGFLDLSLHDQVHLLECAWLEILMIGLV
                     WRSMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEFVCLKS
                     IILLNSGVYTFLSSTLKSLEEKDHIHRVLDKITDTLIHLMAKAGLTLQQQHRRLAQLL
                     LILSHIRHMSNKGMEHLYNMKCKNVVPLYDLLLEMLDAHRLHAPINLGGPPPEDMSQS
                     QLATSGSTPSHSLQMYYITGEAENFPTTI"
BASE COUNT      410 a    530 c    504 g    344 t
ORIGIN

  Query Match             14.2%;  Score 203;  DB 48;  Length 1788;
  Best Local Similarity   67.1%;  Pred. No. 1.71e-174;
  Matches  427;  Conservative    0;  Mismatches 206;  Indels    3;  Gaps    3;

Db   1006 cccttcagtgaggcttcaatgatggacttgctgaccaaccttgcagacagggagctggta 1065
           ||| |||  ||||| || ||||||     |   ||||||||   | || |||| |||| |||||
Qy    703 CCCTTCACCGAGGCCTCCATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTA 762

Db   1066 cacatgatcaactgggcaaagagggtgccaggattttggatttaagcctccatgatcaa 1125
          ||||||||||  ||||| |||| | | || || ||| ||||   | |||||     || |||
Qy    763 CACATGATCAGCTGGGCCAAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAA 822

Db   1126 gtgcatctttcagaatgtgcctggcctagagatcctcatgattggtcttgtctggcgctcc 1185
          ||||  ||  ||||  |     ||| | ||| |   | ||||| || ||   | ||||||||
Qy    823 GTGCGGCTCTTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCA 882

Db   1186 atggagcacccagggaagctcctgtttgctcctaacttgctcctggacaggaaccagggc 1245
          || || ||||| || |||||| | ||||||||  |  |  | |||||||||| |  ||||
Qy    883 ATTGACCACCCCGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGG 942

Db   1246 aagtgtgtcgagggaatggtggagatctttgacatgttgctggctacatcatctcgcttc 1305
```

```
          || || || || |||||  |||| ||||||||||||| | ||||| ||  | ||  | ||
Qy    943 AAATGCGTAGAAGGAATTCTGAAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTT 1002

Db   1306 cgtatgatgaatctccagggagaggagtttgtgtgcctcaaatccatcattttgcttaat 1365
          ||   | | || |||||     | || | | | | ||  ||||  |||| ||  |||| |||
Qy   1003 CGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAAT 1062

Db   1366 tctggagtgtacacatttctgtccagcaccctgaagtctctggaagagaaggaccatatc 1425
          ||  |  ||||| |  |   |   |     |||| | |  ||    | ||  |||   | |
Qy   1063 TCCAGTATGTACCCTCTGGTCACAGCGACCCAGGATGCTGAC-AGCAGCCGGAAGCTGGC 1121

Db   1426 cacagtgtcctggacaagatcacagacactttgatccacctgatggccaaagcaggcctg 1485
              | ||  ||| ||    | || ||  | ||| |      |||| |||||     ||  |
Qy   1122 TCACTTG--CTGAACGCCGTGACCGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATC 1179

Db   1486 actctgcagcagcagcaccggcgtctggcgcagctcctcctcattctatctcacatcagg 1545
            |    |||||||||   |   ||| || ||   | |||||  |   |||||||| |||    ||
Qy   1180 TCCTCCCAGCAGCAATCCATGCGCCTGGCTAACCTCCTGATGCTCCTGTCCCACGTCAGG 1239

Db   1546 cacatgagtaacaaaggcatggagcatctgtacaacatgaagtgcaagaacgtggtgccc 1605
          ||    ||||||||| |||||||| ||||||   |||||||||||||| | | ||||   ||
Qy   1240 CATGCGAGTAACAAGGGCATGGAACATCTGCTCAACATGAAGTGCAAAAATGTGGTCCCA 1299

Db   1606 ctctatgacctgctgctggagatgctggacgcccac 1641
            | |||||||||||||||||||||||| | ||||||
Qy   1300 GTGTATGACCTGCTGCTGGAGATGCTGAATGCCCAC 1335


RESULT     6
LOCUS       OAESTREC     1966 bp    RNA             MAM       08-OCT-1996
DEFINITION  O.aries mRNA for estrogen receptor.
ACCESSION   Z49257
NID         g1617201
KEYWORDS    estrogen receptor.
SOURCE      sheep.
  ORGANISM  Ovis aries
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Artiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Caprinae; Ovis.
REFERENCE   1  (bases 1 to 1966)
  AUTHORS   Madigou,T., Tiffoche,C., Lazennec,G., Pelletier,J. and
            Thieulant,M.L.
  TITLE     The sheep estrogen receptor: cloning and regulation of expression
            in the hypothalamo-pituitary axis
  JOURNAL   Mol. Cell. Endocrinol. 121 (2), 153-163 (1996)
  MEDLINE   97047396
REFERENCE   2  (bases 1 to 1966)
  AUTHORS   Madigou,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-MAY-1995) Thierry Madigou, Interactions Cellulaires
            et Moleculaires, CNRS URA 256, Rennes, 35042, FRANCE
FEATURES             Location/Qualifiers
     source          1..1966
                     /organism="Ovis aries"
                     /strain="Romanov"
                     /clone="lambda P112"
                     /dev_stage="adult"
                     /tissue_type="anterior pituitary gland"
                     /clone_lib="cDNA library"
                     /sex="Female"
     mRNA            1..1966
                     /product="estrogen receptor"
BASE COUNT      439 a    604 c    549 g    374 t
ORIGIN

  Query Match             14.0%;  Score 201;  DB 47;  Length 1966;
  Best Local Similarity   67.0%;  Pred. No. 2.21e-172;
  Matches  426;  Conservative    0;  Mismatches 207;  Indels    3;  Gaps    3;

Db   1158 cctttcagtgaggcttcaatgatgggcttgctgaccggccttgcagacagggagctggta 1217
           || ||||  ||||| || ||||||     |   ||||||      | || |||| |||| |||||
Qy    703 CCCTTCACCGAGGCCTCCATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTA 762
```

```
Db    1218 cacatgatcaactgggcaaagagggtgccaggatttgtggatctggccctccagatcaa 1277
           |||||||||| |||||| |||| | | || || |||||||| ||    |||    ||| |||
Qy     763 CACATGATCAGCTGGGCCAAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAA 822

Db    1278 gtccacctttggaatgtgcctggctagagatcctcatgattggtcttgtctggcgctcc 1337
           || |   || |||||   |     ||| | ||| |   | ||||| || ||   | ||||||||
Qy     823 GTGCGGCTCTTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCA 882

Db    1338 atggagcatccaggaaagctcatctttgctcctaacctcatcctggacaggaacaggga 1397
           || || || || || |||||| | ||||||||   | ||| | ||||||||| |   ||||
Qy     883 ATTGACCACCCGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGG 942

Db    1398 aaatgtgtggaaggcatggtggagatctttgacatgttgctggctacatcgtctcggttc 1457
           ||||| || ||||| ||  |||| |||||||||||| | ||||| ||   | ||  ||||
Qy     943 AAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTT 1002

Db    1458 cgtatgatgaatctccagggagaagagtttgtgtgcctcaaatccatcatcttgcttaat 1517
           ||    | | || |||||       | || | | | ||  ||||  |||| ||| |||| |||
Qy    1003 CGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAAT 1062

Db    1518 tctgggtgtacacatttctgtccagcactctgaggtctctggaagagaaggaccacatc 1577
           ||  |  |||||| |  |  |  | |||   |  |||   |||  ||      |     |
Qy    1063 TCCAGTATGTACCCTCTGGTCAC-AGGGACCC-AGGATGCTGACAGCAGCGGAAGCTGG 1120

Db    1578 cacgcgtcctggacaagatcacagacacctttgatccatctgatggccaagcaggcctg 1637
           | |  | | ||| ||      | || ||  | ||| |       |||| |||||     ||| |
Qy    1121 CTCA-CTTGCTGAACGCCGTGACCGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATC 1179

Db    1638 actctgcagcagcagcaccggcgtctggcccagttcctcctcctcctctcacttcagg 1697
            |     ||||||||    |   ||| |||||  |  ||||  | ||||| || ||| |||||
Qy    1180 TCCTCCCAGCAGCAATCCATGCCCTGGCTAACCTCCTGATGCTCCTGTCCCACGTCAGG 1239

Db    1698 cacatgagcaacaaaggcatggaccacctatacagcatgaagtgcaagaacgtggtgcct 1757
           ||    |||   |||| |||||||| || ||     || ||||||||||||| || ||||| ||
Qy    1240 CATGCGAGTAACAAGGGCATGGAACATCTGCTCAACATGAAGTGCAAAAATGTGGTCCCA 1299

Db    1758 ctctacgacctgctactggagatgctggatgcccac 1793
             | || |||||||||||||||||||||| ||||||||
Qy    1300 CTGTATGACCTGCTCCTGGAGATGCTGAATGCCCAC 1335
```

```
RESULT      7
LOCUS       GGERR        2085 bp    RNA             VRT       12-SEP-1993
DEFINITION  Chicken mRNA for estrogen receptor.
ACCESSION   X03805
NID         g63378
KEYWORDS    DNA binding protein; estrogen receptor; receptor; steroid hormone
            receptor; unidentified reading frame.
SOURCE      chicken.
  ORGANISM  Gallus gallus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Archosauria; Aves; Neognathae; Galliformes;
            Phasianidae; Phasianinae; Gallus
REFERENCE   1  (bases 1 to 2085)
  AUTHORS   Krust,A., Green,S., Argos,P., Kumar,V., Walter,P., Bornert,J.M. and
            Chambon,P.
  TITLE     The chicken oestrogen receptor sequence: homology with v-erbA and
            the human oestrogen and glucocorticoid receptors
  JOURNAL   EMBO J. 5 (5), 891-897 (1986)
  MEDLINE   86247578
COMMENT     Data kindly reviewed (23-JUL-1986) by S. Green.
FEATURES             Location/Qualifiers
     source          1..2085
                     /organism="Gallus gallus"
     mRNA            1..2085
                     /note="ER mRNA"
     CDS             69..131
                     /note="open reading frame 1 (aa 1-20)"
                     /codon_start=1
                     /db_xref="PID:g63379"
                     /translation="MYCCFNMISAEPQNPFWCFF"
     CDS             177..200
                     /note="open reading frame 2 (aa 1-7)"
                     /codon_start=1
                     /db_xref="PID:g584490"
                     /translation="MFLAHCH"
     CDS             203..1972
                     /note="chicken estrogen receptor (aa 1-589)"
                     /codon_start=1
                     /db_xref="PID:g63380"
                     /db_xref="SWISS-PROT:P06212"
                     /translation="MIMILHTKASGVILLHQIQGTELETLSPPQIFIPLERSLSDMYV
                     ESNKTGVFNYPEGATYDFGTTAPVYGSTTLSYAPTSESFGSSSLAGFHSLNNVPPSPV
                     VFLQIAPQLSPFIHHHSQQVPYYLENEQGSFGMREAAPPAFYRPSSDNRRHSIRERMS
                     STNEKGSLSMESTKETRYCAVCNDYASGYHYGVWSCEGCKAFFKRSIQGHNDYMCPAT
                     NQCTIDKNRRKSCQACRLRKCYEVGMMKGGIRKDRRGGEMMKQKRQRFEQDSRNGEAS
                     STELRAPTLWTSPLVVKHNKKNSPALSLTAEQMVSALLEAEPPIVYSEYDPNRPFNEA
                     SMMTLLTNLADRELVHMINWAKRVPGFVDLTLHDQVHLLECAWLEILMIGLVWRSMEH
                     PGKLLFAPNLLLDRNQGKCVEGMVEIFDMLLATAARFRMMNLQGEEFVCLKSIILLNS
                     GVYTFLSSTLKSLEERDYIHRVLDKITDTLIHLMAKSGLSLQQQHRRLAQLLLILSHI
                     RHMSNKGMEHLYNMKCKNVVPLYDLLLEMLDAHRLHAPAARSAAPMEEENRNQLTTAP
                     ASSHSLQSFYINSKEEESMQNTI"
     misc_feature    203..313
                     /note="aa 1-37 -> highly conserved region A"
     misc_feature    314..721
                     /note="conserved region B (aa 38-173)"
     misc_feature    722..970
                     /note="highly conserved region C (aa 174-256) pot.
                     DNA-binding domain"
     misc_feature    971..1087
                     /note="conserved region D (aa 257-295)"
     misc_feature    1088..1840
                     /note="highly conserved region E (aa 296-546) certain
                     estrogen binding domain"
     misc_feature    1841..1969
                     /note="conserved region F (aa 547-589)"
BASE COUNT      582 a    519 c    480 g    504 t
ORIGIN

  Query Match             14.0%;  Score 291;  DB 50;  Length 2085;
  Best Local Similarity   66.7%;  Pred. No. 2.21e-172;
  Matches  431;  Conservative    0;  Mismatches 212;  Indels    2;  Gaps    2;

Db    1180 ccccaatagacccattcaacgaagcatctatgatgaccctgttgaccaaccttgcagacag 1239
             ||||| |    || |||| ||| || || |||||||        |||||||  | || ||||
Qy     693 CCCCAGTGCCCCTTTTACCGAGCCTCCATGATGATGCCCTCAGCAAGTTGCCCGACAA 752

Db    1240 agaattagtgcacatgatcaactgggcaaagagagttccaggatttgtggatttaacact 1299
             || || || |||||||||| |||||| ||||    |||| || ||||||||   | |   ||
Qy     753 GGAGTTGGTACACATGATCAGCTGGGCCAAGAAGATTCCCGGCTTTGTGGAGCTCAGCCT 812

Db    1300 ccatgatcaggtccatctgctggaatgtgcctggttagagatattgatgatcggcttagt 1359
                   || || || |  || ||||  |     ||| | ||| | || ||||| ||  |  |
Qy     813 GTTCGACCAAGTGCGGCTCTTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGAT 872

Db    1360 ctggcgctccatggaacacccaggaaagcttttatttgcacctaatctattactggacag 1419
             |||||||| || || ||||| || |||||  | ||||| ||  |||| | ||||||||
Qy     873 GTGGCGCTCAATTGACCACCCGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAG 932

Db    1420 gaatcaagggaaatgtgtagaaggcatggtggaaatctttgacatgctgctggctactac 1479
           | || | |||||||| ||||| || ||  ||||||||||||||||||||| ||||| ||| |
Qy     933 GGATGAGGGGAAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTAC 992

Db    1480 tgctcggtttcggatgatgaaccttcaagggaggaatttgtgtgccttaaatccatcat 1539
           | |  |||||||     | | || || |||      | |||| | | ||  | ||| |||| ||
Qy     993 TTCAAGGTTTCGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGAT 1052

Db    1540 cctgctcaattctggtgtgtacactttctttctagcaccttgaaatctctggaagagaag 1599
           |||||||||||    || ||||| || |  |  |     ||| | |  ||      |   |
Qy    1053 CCTGCTCAATTCCAGTATGTACCCTCTGGTCACAGGGACCCAGGATGCTGAC-AGCAGCC 1111

Db    1600 ggactatatccaccgtgttctggacaaaatcacagatactctgatccacctaatggcaaa 1659
           |||    |   |    | ||   ||| ||      | || ||| || || |      | || || ||
Qy    1112 GGAAGCTGGCTCACTTG--CTGAACGCCGTGACCGATGCTTTGGTTTGGGTGATTGCCAA 1169

Db    1660 gtcaggtctttctctacagcagcaacaccggcgactaaatcaactcctcttatcctctc 1719
```

```
           |   ||   | ||    |||||||||  |  |||  ||  ||| |  |||||  |   ||||  ||
Qy    1170 GAGCGGCATCTCCTCCCAGCAGCAATCCATGCGCCTGGCTAACCTCCTGATGCTCCTGTC 1229

Db    1720 tcacatcaggcatatgagcaacaaaggaatggagcacctgtacaatatgaagtgtaaaaa 1779
            ||| ||||| || |||| ||||| || ||||| || |||  ||| |||||||| |||||
Qy    1230 CCACGTCAGGTATGTGAGTAACAAGGGCATGGAACATCTGCTCAACATGAAGTGCAAAAA 1289

Db    1780 tgtagttccgctctacgacctcttactggagatgctggacgctcac 1825
           ||| || ||  | || |||||  | |||||||||||| | || |||
Qy    1290 TGTGGTCCCAGTGTATGACCTGCTGCTGGAGATGCTGAATGCCCAC 1335


RESULT     8
LOCUS       I08538       2092 bp                    PAT       24-NOV-1994
DEFINITION  Sequence 1 from Patent WO 8705049
ACCESSION   I08538
NID         g588753
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2092)
  AUTHORS   Shine,J.
  TITLE     EUCARYOTIC EXPRESSION OF STEROID RECEPTOR PROTEINS
  JOURNAL   Patent: WO 8705049 A 1 27-AUG-1987;

FEATURES             Location/Qualifiers
     source          1..2092
                     /organism="unknown"
BASE COUNT      473 a    605 c    593 g    421 t
ORIGIN

  Query Match             13.7%;  Score 197;  DB 55;  Length 2092;
  Best Local Similarity  66.7%;  Pred. No. 3.68e-168;
  Matches   424;  Conservative    0;  Mismatches 209;  Indels   3;  Gaps   2;

Db    1298 cccttcagtgaagcttcgatgatgggcttactgaccaacctggcagacagggagctggtt 1357
           |||||||  || || || ||||||   |  ||||||||  |||| |||| |||| ||||
Qy     703 CCCTTCACCGAGGCCTCCATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTA 762

Db    1358 cacatgatcaactgggcgaagagggtgccaggctttgtggatttgaccctccatgatcag 1417
           |||||||||| |||||| |||| | | || |||||||||||  | | |||    || ||
Qy     763 CACATGATCAGCTGGGCCAAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAA 822

Db    1418 gtccaccttctagaatgtgcctggctagagatcctgatgattggtctcgtctggcgctcc 1477
           || |  ||  | ||  |    ||| | ||| |  | ||||| || ||  | ||||||||
Qy     823 GTGCGGCTCTTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCA 882

Db    1478 atggagcacccagtgaagctactgtttgctcctaacttgctcttggacaggaaccaggga 1537
           || || ||||| |  |||||  | ||||||||  |  |  |  |||||||| |  ||||
Qy     883 ATTGACCACCCCGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGG 942

Db    1538 aaatgtgtagagggcatggtggagatcttcgacatgctgctggctacatcatctcggttc 1597
           ||||| ||||| || ||  |||| ||||| |||||||| ||||| ||  | ||  ||||
Qy     943 AAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTT 1002

Db    1598 cgcatgatgaatctgcagggagaggagtttgtgtgcctcaaatctattattttgcttaat 1657
           ||   | | || || ||     | || | | | ||  ||||  | || ||  |||| |||
Qy    1003 CGAGAGTTAAAACTCCAACATAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAAT 1062

Db    1658 tctggagtgtacacatttctgtccagcaccctgaagtctctggaagagaaggaccatatc 1717
           ||  |  ||||| |  |  |  |    |||| | |  ||    |  ||  |||   |  |
Qy    1063 TCCAGTATGTACCCTCTGGTCACAGCGACCCAGGATGCTGAC-AGCAGCCGGAAGCTGGC 1121

Db    1718 caccgagtcctggacaagatcacagacactttgatccacctgatggccaaggcaggcctg 1777
              |  |  ||| ||    | || ||  ||||| |     |||| ||||||   ||| |
Qy    1122 TCACTTG--CTGAACGCCGTGACCGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATC 1179

Db    1778 accctgcagcagcagcaccagcggctggcccagctcctcctcatcctctcccacatcagg 1837
            ||   ||||||||   |  ||| |||||  | |||||  |  |||| |||||| |||||
Qy    1180 TCCTCCCAGCAGCAATCCATGCGCCTGGCTAACCTCCTGATGCTCCTGTCCCACGTCAGG 1239

Db    1838 cacatgagtaacaaaggcatggagcatctgtacagcatgaagtgcaagaacgtggtgccc 1897
```

```
            |  |||||||||| |||||||| ||||||  || |||||||||||| || ||||| ||
Qy    1240 TATGTGAGTAACAAGGGCATGGAACATCTGCTCAACATGAAGTGCAAAAATGTGGTCCCA 1299

Db    1898 ctctatgacctgctgctggagatgctggacgcccac 1933
            | |||||||||||||||||||||||| | ||||||
Qy    1300 GTGTATGACCTGCTGCTGGAGATGCTGAATGCCCAC 1335


RESULT     9
LOCUS       HUMERMCF     2092 bp    mRNA             PRI       08-NOV-1994
DEFINITION  Human estrogen receptor mRNA, complete cds.
ACCESSION   M12674
NID         g182192
KEYWORDS    estrogen receptor.
SOURCE      Human breast cancer cell line MCF-7, cDNA to mRNA, clone OR8.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2092)
  AUTHORS   Greene,G.L., Gilna,P., Waterfield,M., Baker,A., Hort,Y. and
            Shine,J.
  TITLE     Sequence and expression of human estrogen receptor complementary
            DNA
  JOURNAL   Science 231 (4742), 1150-1154 (1986)
  MEDLINE   86122927
FEATURES             Location/Qualifiers
     source          1..2092
                     /organism="Homo sapiens"
                     /cell_line="MCF-7"
                     /cell_type="Mammary Carcinoma"
                     /tissue_lib="lambda gt11"
                     /lab_host="Prokaryotae"
                     /map="6q24-q27"
     mRNA            <1..>2092
                     /note="ER mRNA"
     CDS             293..2080
                     /gene="ESR"
                     /note="estrogen receptor"
                     /codon_start=1
                     /db_xref="GDB:G00-119-120"
                     /db_xref="PID:g182193"
                     /translation="MTMTLHTKASGMALLHQIQGNELEPLNRPQLKIPLERPLGEVYL
                     DSSKPAVYNYPEGAAYEFNAAAAANAQVYGQTGLPYGPGSEAAAFGSNGLGGFPPLNS
                     VSPSPLMLLHPPPQLSPFLQPHGQQVPYYLENEPSGYTVREAGPPAFYRPNSDNRRQG
                     GRERLASTNDKGSMAMESAKETRYCAVCNDYASGYHYGVWSCEGCKAFFKRSIQGHND
                     YMCPATNQCTIDKNRRKSCQACRLRKCYEVGMMKGGIRKDRRGGRMLKHKRQRDDGEG
                     RGEVGSAGDMRAANLWPSPLMIKRSKKNSLALSLTADQMVSALLDAEPPILYSEYDPT
                     RPFSEASMMGLLTNLADRELVHMINWAKRVPGFVDLTLHDQVHLLECAWLEILMIGLV
                     WRSMEHPVKLLFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEFVCLKS
                     IILLNSGVYTFLSSTLKSLEEKDHIHRVLDKITDTLIHLMAKAGLTLQQQHQRLAQLL
                     LILSHIRHMSNKGMEHLYSMKCKNVVPLYDLLLEMLDAHRLHAPTSRGGASVEETDQS
                     HLATAGSTSSHSLQKYYITGEAEGFPATV"
BASE COUNT      473 a    605 c    593 g    421 t
ORIGIN      1 bp upstream of EcoRI site.

  Query Match             13.7%;  Score 197;  DB 80;  Length 2092;
  Best Local Similarity  66.7%;  Pred. No. 3.68e-168;
  Matches   424;  Conservative    0;  Mismatches 209;  Indels   3;  Gaps   2;

Db    1298 cccttcagtgaagcttcgatgatgggcttactgaccaacctggcagacagggagctggtt 1357
           |||||||  || || || ||||||   |  ||||||||  |||| |||| |||| ||||
Qy     703 CCCTTCACCGAGGCCTCCATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTA 762

Db    1358 cacatgatcaactgggcgaagagggtgccaggctttgtggatttgaccctccatgatcag 1417
           |||||||||| |||||| |||| | | || |||||||||||  | | |||    || ||
Qy     763 CACATGATCAGCTGGGCCAAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAA 822

Db    1418 gtccaccttctagaatgtgcctggctagagatcctgatgattggtctcgtctggcgctcc 1477
           || |  ||  | ||  |    ||| | ||| |  | ||||| || ||  | ||||||||
Qy     823 GTGCGGCTCTTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCA 882

Db    1478 atggagcacccagtgaagctactgtttgctcctaacttgctcttggacaggaaccaggga 1537
           || || ||||| |  |||||  | ||||||||  |  |  |  |||||||| |  ||||
Qy      883 ATTGACCACCCCGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGG 942

Db     1538 aaatgtgtagagggcatggtggagatcttcgacatgctgctggctacatcatctcggttc 1597
            ||||| ||||| || ||  |||| ||||| |||||||| ||||| ||  | ||  ||||
Qy      943 AAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTT 1002

Db     1598 cgcatgatgaatctgcagggagaggagtttgtgtgcctcaaatctattattttgcttaat 1657
            ||   | | || || ||     | || | | | ||  ||||  | || ||  |||| |||
Qy     1003 CGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAAT 1062

Db     1658 tctggagtgtacacatttctgtccagcaccctgaagtctctggaagagaaggaccatatc 1717
            ||  |  ||||| |  |  |  |    |||| | |  ||    |  ||  |||   |  |
Qy     1063 TCCAGTATGTACCCTCTGGTCACAGCGACCCAGGATGCTGAC-AGCAGCCGGAAGCTGGC 1121

Db     1718 caccgagtcctggacaagatcacagacactttgatccacctgatggccaaggcaggcctg 1777
               |  |  ||| ||    | || ||  ||||| |     |||| ||||||   ||| |
Qy     1122 TCACTTG--CTGAACGCCGTGACCGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATC 1179

Db     1778 accctgcagcagcagcaccagcggctggcccagctcctcctcatcctctcccacatcagg 1837
             ||   ||||||||   |  ||| |||||  | |||||  |  |||| |||||| |||||
Qy     1180 TCCTCCCAGCAGCAATCCATGCGCCTGGCTAACCTCCTGATGCTCCTGTCCCACGTCAGG 1239

Db     1838 cacatgagtaacaaaggcatggagcatctgtacagcatgaagtgcaagaacgtggtgccc 1897
            ||   ||||||||| |||||||| ||||||  || |||||||||||| || ||||| ||
Qy     1240 CATGCGAGTAACAAGGGCATGGAACATCTGCTCAACATGAAGTGCAAAAATGTGGTCCCA 1299

Db     1898 ctctatgacctgctgctggagatgctggacgcccac 1933
             | |||||||||||||||||||||||| | ||||||
Qy     1300 GTGTATGACCTGCTGCTGGAGATGCTGAATGCCCAC 1335


RESULT    10
LOCUS       115368          4963 bp    DNA             PAT       09-NOV-1995
DEFINITION  Sequence 16 from patent US 5464758
ACCESSION   115368
NID         g1250276
KEYWORDS    .
SOURCE      Unknown
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 4963)
  AUTHORS   Gossen,M. and Bujard,H.
  TITLE     Tight control of gene expression in eucaryotic cells by
            tetracycline responsive promoters
  JOURNAL   Patent: US 5464758-A 16 07-NOV-1995;

FEATURES             Location/Qualifiers
     source          1..4963
                     /organism="unknown"
BASE COUNT     1259 a   1300 c   1278 g   1126 t
ORIGIN

  Query Match             13.7%;  Score 197;  DB 56;  Length 4963;
  Best Local Similarity   66.7%;  Pred. No  3.68e-168;
  Matches   424;  Conservative    0;  Mismatches 209;  Indels    3;  Gaps    2;

Db     1471 cccttcagtgaagcttcgatgatgggcttactgaccaacctggcagacagggagctggtt 1530
            |||||||  || || || ||||||   |  ||||||||  |||| |||| |||| ||||
Qy      703 CCCTTCACCGAGGCCTCCATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTA 762

Db     1531 cacatgatcaactgggcgaagagggtgccaggctttgtggatttgaccctccatgatcag 1590
            |||||||||| |||||| |||| | | || |||||||||||  | | |||    || ||
Qy      763 CACATGATCAGCTGGGCCAAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAA 822

Db     1591 gtccaccttctagaatgtgcctggctagagatcctgatgattggtctcgtctggcgctcc 1650
            || |  ||  | ||  |    ||| | ||| |  | ||||| || |  |  ||||||||
Qy      823 GTGCGGCTCTTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCA 882

Db     1651 atggagcacccagtgaagctactgtttgctcctaacttgctcttggacaggaaccaggga 1710
            || || ||||| |  |||||  | ||||||||  |  |  |  |||||||| |  ||||
Qy      883 ATTGACCACCCCGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGG 942

Db     1711 aaatgtgtagagggcatggtggagatcttcgacatgctgctggctacatcatctcggttc 1770
            ||||| ||||| || ||  |||| ||||| |||||||| ||||| ||  | ||  ||||
Qy      943 AAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTT 1002

Db     1771 cgcatgatgaatctgcagggagaggagtttgtgtgcctcaaatctattattttgcttaat 1830
            ||   | | || || ||     | || | | | ||  ||||  | || ||  |||| |||
Qy     1003 CGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAAT 1062

Db     1831 tctggagtgtacacatttctgtccagcaccctgaagtctctggaagagaaggaccatatc 1890
            ||  |  ||||| |  |  |  |    |||| | |  ||    |  ||  |||   |  |
Qy     1063 TCCAGTATGTACCCTCTGGTCACAGCGACCCAGGATGCTGAC-AGCAGCCGGAAGCTGGC 1121

Db     1891 caccgagtcctggacaagatcacagacactttgatccacctgatggccaaggcaggcctg 1950
               |  |  ||| ||    | || ||  ||||| |     |||| ||||||   ||| |
Qy     1122 TCACTTG--CTGAACGCCGTGACCGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATC 1179

Db     1951 accctgcagcagcagcaccagcggctggcccagctcctcctcatcctctcccacatcagg 2010
             ||   ||||||||   |  ||| |||||  | |||||  |  |||| |||||| |||||
Qy     1180 TCCTCCCAGCAGCAATCCATGCGCCTGGCTAACCTCCTGATGCTCCTGTCCCACGTCAGG 1239

Db     2011 cacatgagtaacaaaggcatggagcatctgtacagcatgaagtgcaagaacgtggtgccc 2070
            ||   ||||||||| |||||||| ||||||  || |||||||||||| || ||||| ||
Qy     1240 CATGCGAGTAACAAGGGCATGGAACATCTGCTCAACATGAAGTGCAAAAATGTGGTCCCA 1299

Db     2071 ctctatgacctgctgctggagatgctggacgcccac 2106
             | |||||||||||||||||||||||| | ||||||
Qy     1300 GTGTATGACCTGCTGCTGGAGATGCTGAATGCCCAC 1335


RESULT    11
LOCUS       HSERR          6450 bp    RNA             PRI       12-SEP-1993
DEFINITION  Human mRNA for oestrogen receptor.
ACCESSION   X03635 M11457
NID         g31233
KEYWORDS    estrogen receptor; receptor; steroid hormone receptor.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 6450)
  AUTHORS   Green,S., Walter,P., Kumar,V., Krust,A., Bornert,J.M., Argos,P. and
            Chambon,P.
  TITLE     Human oestrogen receptor cDNA: sequence, expression and homology to
            v-erb-A
  JOURNAL   Nature 320 (6058), 134-139 (1986)
  MEDLINE   86146892
COMMENT     Data kindly reviewed (28-OCT-1986) by P. Chambon.
FEATURES             Location/Qualifiers
     source          1..6450
                     /organism="Homo sapiens"
     promoter        26..32
                     /note="CAAT-box"
     promoter        102..109
                     /note="TATA-box"
     misc_RNA        129
                     /note="cap site"
     misc_feature    247..306
                     /note="pot. ORF"
     CDS             361..2148
                     /codon_start=1
                     /product="oestrogen receptor"
                     /db_xref="PID:g31234"
                     /db_xref="SWISS-PROT:P03372"
                     /translation="MTMTLHTKASGMALLHQIQGNELEPLNRPQLKIPLERPLGEVYL
                     DSSKPAVYNYPEGAAYEFNAAAAANAQVYGQTGLPYGPGSEAAAFGSNGLGGFPPLNS
                     VSPSPLMLLHPPPQLSPFLQPHGQQVPYYLENEPSGYTVREAGPPAFYRPNSDNRRQG
                     GRERLASTNDKGSMAMESAKETRYCAVCNDYASGYHYGVWSCEGCKAFFKRSIQGHND
                     YMCPATNQCTIDKNRRKSCQACRLRKCYEVGMMKGGIRKDRRGGRMLKHKRQRDDGEG
                     RGEVGSAGDMRAANLWPSPLMIKRSKKNSLALSLTADQMVSALLDAEPPILYSEYDPT
                     RPFSEASMMGLLTNLADRELVHMINWAKRVPGFVDLTLHDQVHLLECAWLEILMIGLV
                     WRSMEHPVKLLFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEFVCLKS
                     IILLNSGVYTFLSSTLKSLEEKDHIHRVLDKITDTLIHLMAKAGLTLQQQHQRLAQLL
                     LILSHIRHMSNKGMEHLYSMKCKNVVPLYDLLLEMLDAHRLHAPTSRGGASVEETDQS
                     HLATAGSTSSHSLQKYYITGEAEGFPATV"
     misc_feature    442..456
                     /note="pot. nuclear transfer signal"
     misc_feature    1057..1068
                     /note="pot. cAMP dependent kinase phosphorylation site"
     misc_feature    1264..1275
                     /note="pot. cAMP-dependent kinase phosphorylation site"
     misc_feature    4550..4555
                     /note="pot. alternate polyA signal"
     misc_feature    5625..5630
                     /note="pot. alternate polyA signal"
     polyA_signal    6435..6440
     polyA_site      6450
BASE COUNT     1642 a   1496 c   1534 g   1778 t
ORIGIN

  Query Match             13.7%;  Score 197;  DB 74;  Length 6450;
  Best Local Similarity   66.7%;  Pred. No. 3.68e-168;
  Matches  424;  Conservative    0;  Mismatches 209;  Indels    2;  Gaps    2;

Db    1366 cccttcagtgaagcttcgatgatggcttactgaccaacctggcagacagggagctggtt 1425
             |||||  || || || ||||||    |  ||||||||  |||| |||| |||| ||||
Qy     703 CCCTTCACCGAGGCCTCCATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTA 762

Db    1426 cacatgatcaactgggcgaagaggtgccaggctttgtggatttgaccctccatgatcag 1485
              || |||||| |||||| |||| | | || ||||||||||||   | | |||      || ||
Qy     763 CACATGATCAGCTGGGCCAAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAA 822

Db    1486 gtccaccttctagaatgtgcctggctagagatcctgatgattggtctcgtctggcgctcc 1545
             || |   ||   | ||   |      ||| | ||| |   | ||||| || ||   | ||||||||
Qy     823 GTGCGGCTCTTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCA 882

Db    1546 atggagcacccagtgaagctactgtttgctcctaacttgctcttggacaggaaccaggga 1605
             || || ||||| |   |||||   | ||||||||   |   |   |   |||||||| |   ||||
Qy     883 ATTGACTACCCCGGTAAGCTCATCTTTGCTCCAGATCTTGTCCTGGACAGGGATGAGGGG 942

Db    1606 aaatgtgtagagggcatggtggagatcttcgacatgctgctggctacatcatctcggttc 1665
             ||||| ||||| | ||   |||| ||||| |||||||| ||||| ||   | ||   ||||
Qy     943 AAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTT 1002

Db    1666 cgcatgatgaatctgcagggagaggagtttgtgtgcctcaaatctattattttgcttaat 1725
             ||    | | || || ||      | || | | | ||   ||||   | || ||   |||| |||
Qy    1003 CGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAAT 1062

Db    1726 tctggagtgtacacatttctgtccagcaccctgaagtctctggaagagaaggaccatatc 1785
             ||  |   |||||    |   |   |      |||| |  |  ||      |  ||  |||   |  |
Qy    1063 TCCAGTATGTACCCTCTGGTCACAGCCACCCAGGATGCTGAC-AGCAGCCGGAAGCTGGC 1121

Db    1786 caccgagtcctggacaagatcacagacactttgatccacctgatggccaaggcaggcctg 1845
                |   |   ||| ||     |  ||  ||   ||||| |        |||| ||||||     ||| |
Qy    1122 TCACTTG--CTGAACGCCGTGACCGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATC 1179

Db    1846 accctgcagcagcagcaccagcggctggcccagctcctcctcatcctctcccacatcagg 1905
              ||    ||||||||    |   ||| |||||   | |||||   |   |||| |||||| |||||
Qy    1180 TCCTCCCAGCAGCAATCCATGCGCCTGGCTAACCTCCTGATGCTCCTGTCCCACGTCAGG 1239

Db    1906 cacatgagtaacaaaggcatggagcatctgtacagcatgaagtgcaagaacgtggtgccc 1965
             ||     ||||||||| |||||||| ||||||    || |  |||||||||| || ||||| ||
Qy    1240 CATGCGAGTAACAAGGGCATGGAACATCTGCTCAACATGAAGTGCAAAAATGTGGTCCCA 1299

Db    1966 ctctatgacctgctgctggagatgctggacgcccac 2001
               |||||||||||||||||||||||||||| | ||||||
Qy    1300 GTGTATGACCTGCTGCTGGAGATGCTGAATGCCCAC 1335


RESULT   12
LOCUS       ZFIER        2792 bp    mRNA             VRT       23-JUL-1996
DEFINITION  Poephila guttata estrogen receptor mRNA, complete cds.
ACCESSION   L79911
NID         g1449145
KEYWORDS    estrogen receptor.
SOURCE      Poephila guttata female adult hypothalamus-preoptic area cDNA to
            mRNA.
  ORGANISM  Poephila guttata
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Archosauria; Aves; Neognathae; Passeriformes;
            Estrildidae; Estrildinae; Poephila.
REFERENCE   1  (bases 1 to 2792)
  AUTHORS   Jacobs,E.C., Arnold,A.P. and Campagnoni,A.T.
  TITLE     Zebra finch estrogen receptor cDNA: cloning and mRNA expression
  JOURNAL   J. Steroid Biochem. Mol. Biol. 59(2) (1996) In press
FEATURES             Location/Qualifiers
     source          1..2792
                     /organism="Poephila guttata"
                     /dev_stage="adult"
                     /sex="female"
                     /tissue_type="hypothalamus-preoptic area"
     5'UTR           <1..132
     exon            <1..560
                     /note="Exon boundaries based on alignment to chicken ER"
                     /number=1
     misc_feature    127..237
                     /note="domain A"
     CDS             133..1896
                     /codon_start=1
                     /product="estrogen receptor"
                     /db_xref="PID:g1449146"
                     /translation="MTIHTKISGVILLHQIQGIELEILSRPQLKIPLERSLSDMYVET
                     NKTGVFNYPEGATYDFGTTAPVYSSTTLSYAPTSESFGSSSLAGFHSLNSVPPSPVVF
                     LQTAPHWSPFTHHHSQQVPYYLENEQGSFGMREAAPPAFYRPNSDNRPHSIRERMSSA
                     NEKGSLSMESTKETRYCAVCNDYASGYHYGVWSCEGCKAFFKRSIQGHNDYMCPAINQ
                     CTIDKNRRKSCQACRLRKCYEVGMMKGGIRKDRRGGRVMKQKRQREEQDSPNGEASST
                     ELRAPTLWASPLVVKHNKKNSPALSLTAEQMVSALLEAEPPLVYSEYDPNRPFNEASM
                     MTLLTNLADRELVHMINWAKRVPGFVDLTLHDQVHLLECAWLEILMIGLVWRSMEHPG
                     KLLFAPNLLLDRNQGKCVEGMVEIFDMLLATAARFRMMNLQGEEFVCLKSIILLNSGV
                     YTFLSSTLKSLEEKDYIHRVLDKITDTLIHLMAKSGLSLQQQHRRLAQLLLILSHIRH
                     MSNKGMEHLYNMKCKNVVPLYDLLLEMLDAHRLHAPAARSAAPMEEENRSQLTTASAS
                     SHSLQSFYINSKEEENMQNTL"
     misc_feature    238..645
                     /note="domain B"
     exon            561..751
                     /number=2
     misc_feature    646..894
                     /note="domain C, DNA binding domain"
     exon            752..868
                     /number=3
     exon            869..1204
                     /number=4
     misc_feature    895..1011
                     /note="domain D"
     misc_feature    1012..1764
                     /note="domain E, ligand binding domain"
     exon            1205..1343
                     /number=5
     exon            1344..1477
                     /number=6
     exon            1478..1661
                     /number=7
     exon            1662..>2792
                     /number=8
     misc_feature    1765..1893
                     /note="domain F"
     3'UTR           1897..>2792
     polyA_signal    2737..2743
BASE COUNT      794 a    669 c    615 g    714 t
ORIGIN

  Query Match             13.2%;  Score 189;  DB 52;  Length 2792;
  Best Local Similarity   66.0%;  Pred. No. 9.84e-160;
  Matches  420;  Conservative    0;  Mismatches 213;  Indels    3;  Gaps    3;

Db    1114 ccattcaatgaagcttctatgatgaccctgttgaccaacttgcagacagagaattggtg 1173
             || ||||  || || || |||||||       ||||||| | || ||||  || |||||
Qy     703 CCCTTCACCGAGGCCTCCATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTA 762

Db    1174 cacatgatcaactgggcaaagagagttccaggatttgtggatttaacactccatgatcag 1233
             |||||||||| |||||| ||||   |||| || ||||||||  | |  ||     || ||
Qy      763 CACATGATCAGCTGGGCCAAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAA 822

Db     1234 gtccatctactggaatgtgcatggttagagatactgatgatgggcttagtgtggcgctcc 1293
            || |  ||  ||||  |    ||| | ||| |  | ||||||||  |  ||||||||||
Qy      823 GTGCGGCTCTTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCA 882

Db     1294 atggaacacccaggaaagcttttgtttgcacctaacctgttactggacaggaatcaaggg 1353
            || || ||||| || |||||  | ||||| ||  | ||  | ||||||||| || | |||
Qy      883 ATTGACCACCCCGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGG 942

Db     1354 aaatgtgtagagggcatggtggaaatctttgacatgctgctggctactgctgctcggttt 1413
            ||||| ||||| || ||  ||||||||||||||||||| ||||| ||| || |  |||||
Qy      943 AAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTT 1002

Db     1414 cgcatgatgaaccttcaaggggaggaatttgtgtgccttaagtccatcatcctgctcaat 1473
            ||   | | ||  | |||    | |||| | | ||  | ||| |||| ||||||||||||
Qy     1003 CGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAAT 1062

Db     1474 tctggtgtgtacacttttctttccagcaccttgaaatctctggaagagaaagactatatt 1533
            ||  || ||||| || |  |  |    |||  | |  ||    |  ||   ||   |
Qy     1063 TCCAGTATGTACCCTCTGGTCACAGCGACCCAGGATGCTGAC-AGCAGCCGGAAGCTGG- 1120

Db     1534 caccgtgtcctggataaaatcacagatactctgatacacttaatggctaagtcaggtctt 1593
            | |  | | ||| |     | || ||| || || |     | || || |||   ||  |
Qy     1121 CTCACT-TGCTGAACGCCGTGACCGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATC 1179

Db     1594 tctctgcagcagcaacacagacgactggctcagctcctcctcatcctctctcatatcaga 1653
            ||    |||||||||  ||  || |||||| |  ||||  |  |||| || ||  ||||
Qy     1180 TCCTCCCAGCAGCAATCCATGCGCCTGGCTAACCTCCTGATGCTCCTGTCCCACGTCAGG 1239

Db     1654 cacatgagcaacaaaggaatggagcacctgtacaatatgaagtgtaagaatgtagttccg 1713
            ||   ||| ||||| || ||||| || |||  ||| |||||||| || |||||  | ||
Qy     1240 CATGCGAGTAACAAGGGCATGGAACATCTGCTCAACATGAAGTGCAAAAATGTGGTCCCA 1299

Db     1714 ctctacgatctcttgctggagatgctggatgctcac 1749
             | || || ||  |||||||||||||| |||| |||
Qy     1300 GTGTATGACCTGCTGCTGGAGATGCTGAATGCCCAC 1335


RESULT    13
LOCUS       RNESTROR      1880 bp    RNA             ROD       29-OCT-1996
DEFINITION  Rat mRNA for estrogen receptor.
ACCESSION   X61098
NID         g56120
KEYWORDS    estrogen receptor; steroid hormone receptor.
SOURCE      Norway rat
  ORGANISM  Rattus norvegicus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 1880)
  AUTHORS   Maggi,A.M.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-JUN-1991) A.M.A. Maggi, Milano Mol Pharm Lab, Inst
            Pharmacological Sciences, Piazza Durante 11, 20131 Milano, ITALY
REFERENCE   2  (bases 1 to 1880)
  AUTHORS   Spreafico,E., Bettini,E., Pollio,G. and Maggi,A.
  TITLE     Nucleotide sequence of estrogen receptor cDNA from Sprague-Dawley
            rat
  JOURNAL   Eur. J. Pharmacol. 227 (3), 353-356 (1992)
  MEDLINE   93114361
FEATURES             Location/Qualifiers
     source          1..1880
                     /organism="Rattus norvegicus"
                     /strain="Sprague Dawley"
                     /tissue_type="uterus"
     mRNA            1..1880
                     /evidence=experimental
                     /note="estrogen receptor"
     CDS             74..1876
                     /codon_start=1
                     /product="estrogen receptor"
                     /db_xref="PID:g56121"
```

```
                     /db_xref="SWISS-PROT:P06211"
                     /translation="MTMTLHTKASGMALLHQIQGNELEPLNRPQLKMPMERALGEVYV
                     DNSKPAVFNYPEGAAYEFNAAAAAAAAGASAPVYGQSSITYGPGSEAAAFGANSLGAF
                     PQLNSVSPSPLMLLHPPPHVSPFLHPHGHQVPYYLENEPSAYAVRDTGPPAFYRSNSD
                     NRRQNGRERLSSSSEKGNMIMESAKETRYCAVCNDYASGYHYGVWSCEGCKAFFKRSI
                     QGHNDYMCPATNQCTIDKNRRKSCQACRLRKCYEVGMMKGGIRKDRRGGRMLKHKRQR
                     DDLEGRNEMGTSGDMRAANLWPSPLVIKHTKKNSPALSLTADQMVSALLDAEPPLIYS
                     EYDPSRPFSEASMMGLLTNLADRELVHMINWAKRVPGFGDLNLHDQVHLLECAWLEIL
                     MIGLVWRSMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEF
                     VCLKSIILLNSGVYTFLSSTLKSLEEKDHIHRVLDKITDTLIHLMAKAGLTLQQQHRR
                     LAQLLLILSHIRHMSNKGMEHLYNMKCKNVVPLYDLLLEMLDAHRLHAPASRMGVPPE
                     EPSQSQLTTTSSTSAHSLQTYYIPPEAEGFPNTI"
BASE COUNT      438 a    548 c    505 g    389 t
ORIGIN

  Query Match             12.8%;  Score 184;  DB 91;  Length 1880;
  Best Local Similarity   65.7%;  Pred. No. 1.77e-154;
  Matches  417;  Conservative    0;  Mismatches  215;  Indels    3;  Gaps    3;

Db     1094 cccttcagtgaagcctcaatgatgggcttattgaccaacctggcagacagggagctggtt 1153
            |||||||  || ||||| ||||||   |   |||||||  |||| |||| |||| ||||
Qy      703 CCCTTCACGGAGGCCTCCATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTA 762

Db     1154 catatgatcaactgggcaaagagagtgccaggctttggggacttgaatctccacgatcaa 1213
            || ||||||| |||||| |||| | | || ||||||| |||  | |   |   ||| |||
Qy      763 CACATGATCAGCTGGGCCAAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAA 822

Db     1214 gttcaccttctggagtgtgcctggttggagatcctgatgattggtctggtctggcgctcc 1273
            || |  ||  ||||| |    ||| ||||| |  | ||||| || ||| | ||||||||
Qy      823 GTGCGGCTCTTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCA 882

Db     1274 atggaacacccggggaagctcctgtttgctcctaacttgctcttggacaggaatcaaggt 1333
            || || ||||| || |||||| | ||||||||| |  |  |  |||||||| || | ||
Qy      883 ATTGACCACCCCGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGG 942

Db     1334 aaatgtgtagaaggcatggtggagatctttgacatgttactggctacgtcaagtcggttc 1393
            ||||| |||||||| ||  |||| |||||||||||| | ||||| ||  |     ||||
Qy      943 AAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTT 1002

Db     1394 cgcatgatgaacctgcagggagaagagtttgtgtgcctcaaatcaatcattttgcttaat 1453
            ||   | | || || ||     |||| | | | ||  ||||  | || ||  |||| |||
Qy     1003 CGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAAT 1062

Db     1454 tctggagtgtacacatttctatccagcaccttgaagtctctggaagagaaggaccacatc 1513
            ||  |  |||||    |  |  |    |||  | |  ||    |  ||  |||   |
Qy     1063 TCCAGTATGTACCCTCTGGTCACAGCGACCCAGGATGCTGAC-AGCAGCCGGAAG-CTGG 1120

Db     1514 caccgagtcctggacaagatcacagacactttgatccacttgatagccaaagctggcctg 1573
            | |    | ||| ||    | || ||  ||||| |     |||| |||||    ||| |
Qy     1121 CTCACT-TGCTGAACGCCGTGACCGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATC 1179

Db     1574 actctgcagcaacagcatcagcgtctggcccagctcctcctcatcctttcccatatccgg 1633
             |    ||||| ||      ||| |||||  | |||||  |  |||| |||||  || ||
Qy     1180 TCCTCCCAGCAGCAATCCATGCGCCTGGCTAACCTCCTGATGCTCCTGTCCCACGTCAGG 1239

Db     1634 cacatgagtaacaaaggcatggagcatctctacaacatgaaatgcaagaatgtcgtgcct 1693
            ||   ||||||||| |||||||| |||||   ||||||||| ||||| ||||| || ||
Qy     1240 CATGCGAGTAACAAGGGCATGGAACATCTGCTCAACATGAAGTGCAAAAATGTGGTCCCA 1299

Db     1694 ctctatgacctgctgctggagatgctggatgctca 1728
             | |||||||||||||||||||||||| |||| ||
Qy     1300 GTGTATGACCTGCTGCTGGAGATGCTGAATGCCCA 1334


RESULT    14
LOCUS       RNERR         2090 bp    RNA             ROD       12-SEP-1993
DEFINITION  Rat mRNA for estrogen receptor.
ACCESSION   Y00102
NID         g56110
KEYWORDS    estrogen receptor.
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 2090)
  AUTHORS   Muramatsu,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-MAR-1987) Masami Muramatsu,Dept of Biochemistry,
            Faculty of Medicine The University of Tokyo  Hongo 7-3-1,
            Bunkyo-Ku, Tokyo 113 Japan
REFERENCE   2  (bases 1 to 2090)
  AUTHORS   Koike,S., Sakai,M. and Muramatsu,M.
  TITLE     Molecular cloning and characterization of rat estrogen receptor
            cDNA
  JOURNAL   Nucleic Acids Res. 15 (6), 2499-2513 (1987)
  MEDLINE   87174780
FEATURES             Location/Qualifiers
     source          1..2090
                     /organism="Rattus norvegicus"
                     /strain="Wistar"
                     /tissue_type="utems"
                     /clone_lib="rat uterus cDNA 1 "
     CDS             210..2012
                     /note="estrogen receptor protein (AA 1-600)"
                     /codon_start=1
                     /db_xref="PID:g56111"
                     /db_xref="SWISS-PROT:P06211"
                     /translation="MTMTLHTKASGMALLHQIQGNELEPLNRPQLKMPMERALGEVYV
                     DNSKPAVFNYPEGAAYEFNAAAAAAAAGASAPVYGQSSITYGPGSEAAAFGANSLGAF
                     PQLNSVSPSPLMLLHPPPHVSPFLHPHGHQVPYYLENEPSAYAVRDTGPPAFYRSNSD
                     NRRQNGRERLSSSSEKGNMIMESAKETRYCAVCNDYASGYHYGVWSCEGCKAFFKRSI
                     QGHNDYMCPATNQCTIDKNRRKSCQACRLRKCYEVGMMKGGIRKDRRGGRMLKHKRQR
                     DDLEGRNEMGTSGDMRAANLWPSPLVIKHTKKNSPALSLTADQMVSALLDAEPPLIYS
                     EYDPSRPFSEASMMGLLTNLADRELVHMINWAKRVPGFGDLNLHDQVHLLECAWLEIL
                     MIGLVWRSMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEF
                     VCLKSIILLNSGVYTFLSSTLKSLEEKDHIHRVLDKINDTLIHLMAKAGLTLQQQHRR
                     LAQLLLILSHIRHMSNKGMEHLYNMKCKNVVPLYDLLLEMLDAHRLHAPASRMGVPPE
                     EPSQSQLTTTSSTSAHSLQTYYIPPEAEGFPNTI"
BASE COUNT      488 a    619 c    537 g    446 t
ORIGIN

  Query Match             12.8%;  Score 184;  DB 91;  Length 2090;
  Best Local Similarity   65.7%;  Pred. No. 1.77e-154;
  Matches  417;  Conservative    0;  Mismatches  215;  Indels    3;  Gaps    3;

Db    1230 cccttcagtgaagcctcaatgatgggcttattgaccaacctggcagacagggagctggtt 1289
           |||||||  || ||||| ||||||     |    |||||||  |||| |||| |||| ||||
Qy     703 CCCTTCACCGAGGCCTCCATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTA 762

Db    1290 catatgatcaactgggcaaagagagtgccaggctttggggacttgaatctccatgatcaa 1349
           || |||||||| |||||| ||||    | || ||||||| |||  | |   ||   ||| |||
Qy     763 CACATGATCAGCTGGGCCAAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAA 822

Db    1350 gttcaccttctggagtgtgcctggttggagatcctgatgattggtctcgtctggcgctcc 1409
           || |  ||  ||||| |    ||| ||||| |  | ||||| || ||| | ||||||||
Qy     823 GTGCGGTGCTTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCA 882

Db    1410 atggaacacccggggaagctcctgtttgctcctaacttgctcttggacaggaatcaaggt 1469
           || || ||||| || |||||| | ||||||||   |   |   |||||||| || | ||
Qy     883 ATTGACCACCCCGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGG 942

Db    1470 aaatgtgtagaaggcatggtggagatctttgacatgttgcttgctacgtcaagtcgattc 1529
           ||||| |||||||| ||  |||| ||||||||||||| | |||||| ||   |      | ||
Qy     943 AAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTT 1002

Db    1530 cgcatgatgaacctgcagggagaagagtttgtgtgcctcaaatcaatcattttacttaat 1589
           ||   | | || || ||      |||| | | | ||  ||||  | || ||  |||| |||
Qy    1003 CGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAAT 1062

Db    1590 tctggagtgtacacatttctgtccagcaccttgaagtctctggaagagaaggaccacatc 1649
           ||  |  ||||| |  |  |  |     ||| |  | ||     |  ||  |||    |
Qy    1063 TCCAGTATGTACCCTCTGGTCACAGCGACCCAGGATGCTGAC-AGCAGCCGGAAG-CTGG 1120

Db    1650 caccgagtcctggacaagatcacagacactttgatccacttgatggccaaagctggcctg 1709
             |     | ||| ||     | | |||  ||||| |      |||| |||||     ||| |
```

```
Qy    1121 CTCACT-TGCTGAAGGCGGTGACGGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATC 1179

Db    1710 actctgcagcagcagcatcgccgcctagctcagctccttctcattctttcccatatccgg 1769
            |    ||||| ||        || ||||  | |||||  |  ||||  ||||  || ||
Qy    1180 TCCTCCCAGCAGCAATCCATGCGGCTGGCTAACCTCCTGATGCTCCTGTCCCACGTCAGG 1239

Db    1770 cacatgagtaacaaaggcatggagcatctctacaacatgaaatgcaagaatgtcgtgccc 1829
           ||    ||||||||| ||||||||| |||||    ||||||||| ||||| ||||| || ||
Qy    1240 CATGCGAGTAACAAGGGCATGGAACATCTGCTCAACATGAAGTGCAAAAATGTGGTCCCA 1299

Db    1830 ctctatgacctgctgctggagatgctggatgctca 1864
            | ||||||||||||||||||||||||  |||| ||
Qy    1300 GTGTATGACCTGCTGCTGGAGATGCTGAATGCCCA 1334


RESULT   15
LOCUS       S52128       2090 bp    mRNA            ROD       12-MAR-1993
DEFINITION  estrogen receptor [rats, brain, mRNA, 2090 nt].
ACCESSION   S52128
NID         g262708
KEYWORDS    .
SOURCE      Rattus sp. brain.
  ORGANISM  Rattus sp.
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 2090)
  AUTHORS   Bettini,E., Pollio,G., Santagati,S. and Maggi,A.
  TITLE     Estrogen receptor in rat brain: presence in the hippocampal
            formation
  JOURNAL   Neuroendocrinology 56 (4), 502-508 (1992)
  MEDLINE   93116840
  REMARK    GenBank staff at the National Library of Medicine created this
            entry [NCBI gibbsq 111564] from the original journal article.
            This sequence comes from Fig. 1.
FEATURES             Location/Qualifiers
     source          1..2090
                     /organism="Rattus sp."
BASE COUNT      488 a    619 c    537 g    446 t
ORIGIN

  Query Match             12.8%;  Score 184;  DB 92;  Length 2090;
  Best Local Similarity   65.7%;  Pred. No. 1.77e-154;
  Matches  417;  Conservative    0;  Mismatches  215;  Indels    3;  Gaps    3;

Db    1230 cccttcagtgaagcctcaatgatgggcttattgaccaacctggcagacagggagctggtt 1289
           |||||||  || ||||| ||||||     |    |||||||  |||| |||| |||| ||||
Qy     703 CCCTTCACCGAGGCCTCCATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTA 762

Db    1290 catatgatcaactgggcaaagagagtgccaggctttggggacttgaatctccatgatcaa 1349
           || |||||||| |||||| ||||    | || ||||||| |||  | |   ||   ||| |||
Qy     763 CACATGATCAGCTGGGCCAAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAA 822

Db    1350 gttcaccttctggagtgtgcctggttggagatcctgatgattggtctcgtctggcgctcc 1409
           || |  ||  ||||| |    ||| ||||| |  | ||||| || ||| | ||||||||
Qy     823 GTGCGGTGCTTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCA 882

Db    1410 atggaacacccggggaagctcctgtttgctcctaacttgctcttggacaggaatcaaggt 1469
           || || ||||| || |||||| | ||||||||   |   |   |||||||| || | ||
Qy     883 ATTGACCACCCCGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGG 942

Db    1470 aaatgtgtagaaggcatggtggagatctttgacatgttgcttgctacgtcaagtcgattc 1529
           ||||| |||||||| ||  |||| ||||||||||||| | |||||| ||   |      | ||
Qy     943 AAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTT 1002

Db    1530 cgcatgatgaacctgcagggagaagagtttgtgtgcctcaaatcaatcattttacttaat 1589
           ||   | | || || ||      |||| | | | ||  ||||  | || ||  |||| |||
Qy    1003 CGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAAT 1062

Db    1590 tctggagtgtacacatttctgtccagcaccttgaagtctctggaagagaaggaccacatc 1649
           ||  |  ||||| |  |  |  |     ||| |  | ||     |  ||  |||    |
Qy    1063 TCCAGTATGTACCCTCTGGTCACAGCGACCCAGGATGCTGAC-AGCAGCCGGAAG-CTGG 1120
```

```
Db   1650 caccgagtcctggacaagatcaacgacactttgatccacttgatggccaaagctggcctg 1709
          | |    | ||| ||    | | |||  ||||| |     |||| |||||    ||| |
Qy   1121 CTCACT-TGCTGAACGCCGTGACCGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATC 1179

Db   1710 actctgcagcaacagcatcgccgtctggcccagctcctcctcatcctttcccatatccgg 1769
           |    ||||| ||       || |||||  | |||||  |  |||| |||||  || ||
Qy   1180 TCCTCCCAGCAGCAATCCATGCGCCTGGCTAACCTCCTGATGCTCCTGTCCCACGTCAGG 1239

Db   1770 cacatgagtaacaaaggcatggagcatctctacaacatgaaatgcaagaatgtcgtgcct 1829
          ||   ||||||||| |||||||| |||||   ||||||||| ||||| ||||| || ||
Qy   1240 CATGCGAGTAACAAGGGCATGGAACATCTGCTCAACATGAAGTGCAAAAATGTGGTCCCA 1299

Db   1830 ctctatgacctgctgctggagatgctggatgctca 1864
           | |||||||||||||||||||||||| |||| ||
Qy   1300 GTGTATGACCTGCTGCTGGAGATGCTGAATGCCCA 1334

Search completed: Thu Dec 18 17:57:27 1997
Job time : 2331 secs.
```

```
****************************************************************

                      MPSRCH   (TM)

****************************************************************

         Release 2 1D John F  Collins, Biocomputing Research Unit.
         Copyright (c) 1993, 1994, 1995  University of Edinburgh, U K.
                 Distribution rights by IntelliGenetics, Inc.

MPsrch_nn   n.a. - n.a. database search, using Smith-Waterman algorithm

Run on            Thu Dec 18 17:57:46 1997;  MasPar time 159.60 Seconds
                                             937.746 Million cell updates/sec

Tabular output not generated.

Title:            >US-08-826-361A-1
Description:      (1-1434) from US08826361A seq
Perfect Score     1434
N.A. Sequence.          1 ATGAATTACAGCATTCCCAG    .    .AGAACCCACAGTCTCAGTGA 1434
    Comp.                 TACTTAATGTCGTAAGGGTC    ..   .TCTTGGGTGTCAGAGTCACT

Scoring table:    TABLE default
                  Gap 6

Nmatch   STD :    Dbase 0;  Query 0

Searched.         142090 seqs, 52183452 bases x 2

Post-processing:  Minimum Match 0%
                  Listing first 45 summaries

Database:         n-geneseq28
                    1 part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
                    8:part8 9:part9 10 part10 11:part11 12:part12 13:part13
                    14:part14 15:part15 16:part16 17:part17 18:part18
                    19:part19 20:part20 21:part21 22:part22 23:part23
                    24:part24 25:part25 26:part26 27:part27 28:part28
                    29:part29

Statistics:       Mean 9.341;  Variance 5.204;  scale 1.795

     Pred. No  is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                   SUMMARIES

                       %
Result              Query
  No     Score      Match Length DB   ID              Description              Pred. No.
----------------------------------------------------------------------------------------
    1     1432       99.9   1460 29  T62843          Human oestrogen recep   0.00e+00
    2     1012       70.6   2569 29  T62842          Rat oestrogen recepto   0.00e+00
    3      992       69.2   1458 29  T62844          Mouse oestrogen recep   0.00e+00
    4      197       13.7   4963 17  T06873          PhCMV*-1 promoter and   1.14e-121
    5      197       13.7   4963 13  Q76270          PhCMV*-1-controlled p   1.14e-121
    6      195       13.6   2042  2  N70990          cDNA encoding human o   3.51e-120
    7      127        8.9    204 14  Q85976          Human estrogen recept   4.22e-70
    8       91        6.3   1047  2  Q10572          Human Natriuretic Pep   2.62e-44
c   9       71        5.0   1047  2  Q10572          Human Natriuretic Pep   2.10e-30
c  10       71        5.0   3460 24  T42192          Fragment of pig oestr   2.10e-30
   11       68        4.7   2153  2  N80922          Sequence encoding hum   2.33e-28
   12       63        4.4   6244 13  Q76264          PhCMV*-1-controlled p   5.58e-25
   13       63        4.4   6244 17  T06872          PhCMV*-1 promoter and   5.58e-25
   14       59        4.1   2992 13  Q79434          Human liver hap cDNA    2.63e-22
   15       59        4.1   2992 11  Q65571          Human liver hap cDNA    2.63e-22
   16       59        4.1   2992 20  T06491          Human hepatoma retino   2.63e-22
   17       58        4.0   1912  2  Q10388          Encodes skin-specific   1.21e-21
   18       58        4.0   2521  3  Q20542          Mouse retinoic acid r   1.21e-21
   19       58        4.0   2740  3  Q20541          Mouse retinoic acid r   1.21e-21
   20       58        4.0   5040  2  N80916          Sequence encoding the   1.21e-21
   21       56        3.9   1913 18  T05202          Murine retinoic acid    2.54e-20
   22       56        3.9   6134 17  T49216          Plasmid pGR403F for     2.54e-20
   23       55        3.8   2989  1  N90093          Human Hap (hepatoma)    1.16e-19
   24       53        3.7   1231  6  Q37761          Sequence of a 1.23 kb   2.35e-18
   25       53        3.7   1785  2  Q12005          Human TR2-9 DNA bindi   2.35e-18
   26       53        3.7   2029  2  Q12004          Human TR2-5 androgen    2.35e-18
   27       53        3.7   2221  2  Q12006          Human TR2-11 DNA bind   2.35e-18
   28       53        3.7   2458  2  Q12003          Human TR2-7 DNA bindi   2.35e-18
   29       51        3.6   1576  2  Q10226          Clone pGEM-hRARgamma    4.68e-17
   30       51        3.6   1903  2  Q10389          Encodes human Retinoi   4.68e-17
   31       51        3.6   1903 18  T05196          Human retinoic acid r   4.68e-17
   32       51        3.6   3569  1  N91772          Human androgen recept   4.68e-17
   33       51        3.6   3569 29  T63407          Androgen receptor cDN   4.68e-17
   34       51        3.6   3715  2  Q12001          Full-length human and   4.68e-17
   35       50        3.5   1351 13  Q79935          Human liver RAR-beta    2.07e-16
   36       50        3.5   2249  2  N80921          Sequence encoding hum   2.07e-16
c  37       49        3.4    204  1  N81164          Base substituted E.co   9.09e-16
   38       47        3.3    387 19  T09863          Modified human glucoc   1.72e-14
   39       47        3.3    387 24  T45454          Modified human glucoc   1.72e-14
   40       47        3.3   3217  1  N91578          Rat androgen receptor   1.72e-14
   41       45        3.1   1895  2  Q12008          TrpE/androgen recepto   3.18e-13
   42       45        3.1   3217  2  Q12002          Full-length rat andro   3.18e-13
   43       45        3.1   3427 10  Q56733          cDNA sequence encodin   3.18e-13
   44       45        3.1   4180  1  N91773          Rat androgen receptor   3.18e-13
   45       44        3.1   4400 18  T16151          Apoptopic cerebral ne   1.35e-12

                                  ALIGNMENTS

RESULT     1
ID   T62843 standard; cDNA; 1460 BP.
AC   T62843;
DT   08-JUN-1997  (first entry)
DE   Human oestrogen receptor beta (ER-beta) cDNA.
KW   Orphan receptor; oestrogen receptor beta; ER-beta;
KW   nuclear receptor; prostate cancer; benign prostatic hyperplasia;
KW   ovary cancer; cardiovascular disease; osteoporosis,
KW   environment; pollutant; ss.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   CDS             3..1460
FT   /*tag=  a
PN   WO9709348-A2.
PD   13-MAR-1997.
PF   09-SEP-1996; E03933.
PR   08-SEP-1995; GB-018272.
PR   15-MAR-1996; GB-005550.
PR   11-APR-1996; GB-007532.
PR   08-MAY-1996; GB-009576.
PA   (KARO-) KARO BIO AB.
PI   Enmark E,  Gustafsson J,  Kuiper GG;
DR   WPI; 97-192842/17
DR   P-PSDB, W14724.
PT   New isolated oestrogen receptor beta - used to develop prods  for
PT   treating e.g. cancers, CNS diseases, osteoporosis or cardiovascular
PT   disease
PS   Claim 6; Fig 13B; 45pp, English.
CC   A cDNA clone (T62843) codes for a novel human oestrogen receptor-
CC   related nuclear receptor (W13724), designated ER-beta.  It was
CC   isolated from a human testis cDNA library using rat ER-beta cDNA
CC   (see also T62842) as probe.  The human ER-beta gene was mapped
CC   to chromosome 14 region 14q22-23.  Rat, human and mouse ER-beta
CC   (W14723-25) can be used to isolate molecules for use in the
CC   treatment of cardiovascular diseases, central nervous system
CC   diseases, osteoporosis, prostate or ovarian cancer or benign
CC   prostatic hyperplasia and to test environmental chemicals for
CC   oestrogenic activity.
SQ   Sequence   1460 BP;    352 A;    379 C;    411 G;    318 T;
```

```
  Query Match             99.9%;  Score 1432;  DB 29;  Length 1460;
  Best Local Similarity   99.9%;  Pred. No. 0.00e+00;
  Matches 1433;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Db      27 atgaattacagcattcccagcaatgtcactaacttggaaggtgggcctggtcggcagacc 86
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1 ATGAATTACAGCATTCCCAGCAATGTCACTAACTTGGAAGGTGGGCCTGGTCGGCAGACC 60

Db      87 acaagcccaaatgtgttgtggccaacacctgggcacctttctcctttagtggtccatcgc 146
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      61 ACAAGCCCAAATGTGTTGTGGCCAACACCTGGGCACCTTTCTCCTTTAGTGGTCCATCGC 120

Db     147 cagttatcacatctgtatgcggaacctcaaaagagtccctggtgtgaagcaagatcgcta 206
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     121 CAGTTATCACATCTGTATGCGGAACCTCAAAAGAGTCCCTGGTGTGAAGCAAGATCGCTA 180

Db     207 gaacacaccttacctgtaaacagagagacactgaaaaggaaggttagtgggaaccgttgc 266
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     181 GAACACACCTTACCTGTAAACAGAGAGACACTGAAAAGGAAGGTTAGTGGGAACCGTTGC 240

Db     267 gccagccctgttactggtccaggttcaaagagggatgctcacttctgcgctgtctgcagc 326
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     241 GCCAGCCCTGTTACTGGTCCAGGTTCAAAGAGGGATGCTCACTTCTGCGCTGTCTGCAGC 300

Db     327 gattacgcatcgggatatcactatggagtctggtcgtgtgaaggatgtaaggcctttttt 386
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     301 GATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATGTAAGGCCTTTTTT 360

Db     387 aaaagaagcattcaaggacataatgattatatttgtccagctacaaatcagtgtacaatc 446
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     361 AAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCAGTGTACAATC 420

Db     447 gataaaaaccggcgcaagagctgccaggcctgccgacttcggaagtgttacgaagtggga 506
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     421 GATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTACGAAGTGGGA 480

Db     507 atggtgaagtgtggctcccggagagagagatgtgggtaccgccttgtgcggagacagaga 566
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     481 ATGGTGAAGTGTGGCTCCCGGAGAGAGAGATGTGGGTACCGCCTTGTGCGGAGACAGAGA 540

Db     567 agtgccgacgagcagctgcactgtgccggcaaggccaagagaagtggcggccacgcgccc 626
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     541 AGTGCCGACGAGCAGCTGCACTGTGCCGGCAAGGCCAAGAGAAGTGGCGGCCACGCGCCC 600

Db     627 cgagtgcgggagctgctgctggacgccctgagccccgagcagctagtgctcaccctcctg 686
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     601 CGAGTGCGGGAGCTGCTGCTGGACGCCCTGAGCCCCGAGCAGCTAGTGCTCACCCTCCTG 660

Db     687 gaggctgagccgccccatgtgctgatcagccgccccagtgcgcccttcaccgaggcctcc 746
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     661 GAGGCTGAGCCGCCCCATGTGCTGATCAGCCGCCCCAGTGCGCCCTTCACCGAGGCCTCC 720

Db     747 atgatgatgtccctgaccaagttggccgacaaggagttggtacacatgatcagctgggcc 806
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     721 ATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTACACATGATCAGCTGGGCC 780

Db     807 aagaagattcccggctttgtggagctcagcctgttcgaccaagtgcggctcttggagagc 866
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     781 AAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAAGTGCGGCTCTTGGAGAGC 840

Db     867 tgttggatggaggtgttaatgatggggctgatgtggcgctcaattgaccaccccggcaag 926
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     841 TGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCAATTGACCACCCCGGCAAG 900

Db     927 ctcatctttgctccagatcttgttctggacagggatgaggggaaatgcgtagaaggaatt 986
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     901 CTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGGAAATGCGTAGAAGGAATT 960

Db     987 ctggaaatctttgacatgctcctggcaactacttcaaggtttcgagagttaaaactccaa 1046
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     961 CTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTTCGAGAGTTAAAACTCCAA 1020

Db    1047 cacaaagaatatctctgtgtcaaggccatgatcctgctcaattccagtatgtaccctctg 1106
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    1021 CACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAATTCCAGTATGTACCCTCTG 1080

Db    1107 gtcacagcgacccaggatgctgacagcagccggaagctggctcacttgctgaacgccgtg 1166
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    1081 GTCACAGCGACCCAGGATGCTGACAGCAGCCGGAAGCTGGCTCACTTGCTGAACGCCGTG 1140

Db    1167 accgatgctttggtttgggtgattgccaagagcggcatctcctcccagcagcaatccatg 1226
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    1141 ACCGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATCTCCTCCCAGCAGCAATCCATG 1200

Db    1227 cgcctggctaacctcctgatgctcctgtcccacgtcaggcatgcgagtaacaagggcatg 1286
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    1201 CGCCTGGCTAACCTCCTGATGCTCCTGTCCCACGTCAGGCATGCGAGTAACAAGGGCATG 1260

Db    1287 gaacatctgctcaacatgaagtgcaaaaatgtggtcccagtgtatgacctgctgctggag 1346
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    1261 GAACATCTGCTCAACATGAAGTGCAAAAATGTGGTCCCAGTGTATGACCTGCTGCTGGAG 1320

Db    1347 atgctgaatgcccacgtgcttcgcgggtgcaagtcctccatcacggggtccgagtgcagc 1406
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    1321 ATGCTGAATGCCCACGTGCTTCGCGGGTGCAAGTCCTCCATCACGGGGTCCGAGTGCAGC 1380

Db    1407 ccggcagaggacagtaaaagcaaagagggctcccagaacctacagtctcagtga 1460
           |||||||||||||||||||||||||||||||||||||||| |||||||||||||
Qy    1381 CCGGCAGAGGACAGTAAAAGCAAAGAGGGCTCCCAGAACCCACAGTCTCAGTGA 1434


RESULT    2
ID   T62842 standard; cDNA; 2568 BP.
AC   T62842;
DT   08-JUN-1997  (first entry)
DE   Rat oestrogen receptor beta (ER-beta) cDNA.
KW   Orphan receptor; oestrogen receptor beta; ER beta;
KW   nuclear receptor; prostate cancer; benign prostatic hyperplasia;
KW   ovary cancer; cardiovascular disease; osteoporosis;
KW   environment; pollutant; ss.
OS   Rattus sp.
FH   Key              Location/Qualifiers
FT   CDS              424..1881
FT   /*tag=  a
FT   /note= "alternative start codon at 448..450"
PN   WO9709348-A2.
PD   13-MAR-1997.
PF   09-SEP-1996; E03933.
PR   08-SEP-1995; GB-018272.
PR   15-MAR-1996; GB-005550.
PR   11-APR-1996; GB-007532.
PR   08-MAY-1996; GB-009576.
PA   (KARO-) KARO BIO AB.
PI   Enmark E,  Gustafsson J,  Kuiper GG;
DR   WPI; 97-192842/17.
DR   P-PSDB; W14723.
PT   New isolated oestrogen receptor beta - used to develop prods. for
PT   treating e.g. cancers, CNS diseases, osteoporosis or cardiovascular
PT   disease
PS   Claim 6; Fig 1; 45pp; English.
CC   A cDNA clone (T62842) codes for a novel rat oestrogen receptor-
CC   related nuclear receptor (W14723), designated ER-beta.  A set of
CC   degenerate primers based on conserved sequences of the DNA-
CC   binding domain and ligand binding domain of nuclear receptors were
CC   used to amplify rat prostate cDNA.  The PCR prod. was used to
CC   screen a rat prostate cDNA library to obtain the ER-beta cDNA.
CC   This was then used to obtain human ER-beta cDNA (see also T62843)
CC   from a human testis cDNA library.  Rat, human and mouse ER-beta
CC   (W14723-25) can be used to isolate molecules for use in the
CC   treatment of cardiovascular diseases, central nervous system
CC   diseases, osteoporosis, prostate or ovarian cancer or benign
CC   prostatic hyperplasia and to test environmental chemicals for
CC   oestrogenic activity.
SQ   Sequence   2568 BP;    734 A;    664 C;    553 G;    617 T;
```

```
  Query Match             70.6%;  Score 1012;  DB 29;  Length 2568;
  Best Local Similarity   85.3%;  Pred. No. 0.00e+00;
  Matches 1223;  Conservative    0;  Mismatches 211;  Indels    0;  Gaps    0;

Db        448 atgaactacagtgttcccggcagcaccagtaacctggacggtgggcctgtccgactgagc 507
              | ||| |||||   ||||| |||      || ||||  |||| ||||||||||   || | || |
Qy          1 ATGAATTACAGCATTCCCAGCAATGTCACTAACTTGGAAGGTGGGCCTGGTCGGCAGACC 60

Db        508 acaagcccaaatgtgctatggccaacttctgggcacctgtctcctttagcgacccattgc 567
              ||||||||||||||| | ||||||||  |||||||||| |||||||||| |  |||| ||
Qy         61 ACAAGCCCAAATGTGTTGTGGCCAACACCTGGGCACCTTTCTCCTTTAGTGGTCCATCGC 120

Db        568 caatcatcgctcctctatgcagaacctcaaaagagtccttggtgtgaagcaagatcacta 627
              |  | | | |   || ||||| |||||||||||||||| |||||||||||||||| |||
Qy        121 CAGTTATCACATCTGTATGCGGAACCTCAAAAGAGTCCCTGGTGTGAAGCAAGATCGCTA 180

Db        628 gagcacaccttacctgtaaacagagagacactgaagaggaagcttagtgggagcagttgt 687
              || |||||||||||||||||||||||||||||||| ||||| ||||||||| | ||||
Qy        181 GAACACACCTTACCTGTAAACAGAGAGACACTGAAAAGGAAGGTTAGTGGCAACCGTTGC 240

Db        688 gccagccctgttactagtccaaacgcaaagagggatgctcacttctgccccgtctgcagc 747
              ||||||||||||||| ||||     ||||||||||||||||||||||||| | |||||||||
Qy        241 GCCAGCCCTGTTACTGGTCCAGGTTCAAAGAGGGATGCTCACTTCTGCGCTGTCTGCAGC 300

Db        748 gattatgcatctgggtatcattacggcgtttggtcatgtgaaggatgtaaggcctttttt 807
              ||||| |||||  || ||||| ||  ||  || |||||  ||||||||||||||||||||||||
Qy        301 GATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATGTAAGGCCTTTTTT 360

Db        808 aaaagaagcattcaaggacataatgattatatctgtccagccacgaatcagtgtaccata 867
              |||||||||||||||||||||||||||||||| ||||||||| || |||||||||| ||
Qy        361 AAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCAGTGTACAATC 420

Db        868 gacaagaaccggcgtaaaagctgccaggcctgccgacttcgcaagtgttatgaagtagga 927
              || || ||||||||| || |||||||||||||||||||||||| |||||||| ||||| |||
Qy        421 GATAAAAACCGGCGCAAGAGTTGCCAGGCCTGCCGACTTCGGAAGTGTTACGAAGTGGGA 480

Db        928 atggtcaagtgtggatccaggagagaacggtgtgggtaccgtatagtgcggaggcagaga 987
              ||||| ||||||||| ||| |||||||   | ||||||||||| | ||||||||| ||||||
Qy        481 ATGGTGAAGTGTGGCTCCCGGAGAGAGAGATGTGGGTACCGCCTTGTGCGGAGACAGAGA 540

Db        988 agttctagcgaacagtataactgcctgagcaaggccaagagaaacggtgggcatgcaccc 1047
              ||| |    ||||||| | |||||       |||| ||||||||||  || || || || |||
Qy        541 AGTGCCGACGAGCAGCTGCACTGTGCCGGCAAGGCCAAGAGAAGTGGCGGCCACGCGCCC 600

Db       1048 cgggtgaaggagctactgctgagcaccttgagtccagagcaactggtgctcaccctcctg 1107
              || |||  ||||| |||||  |  || |||| || |||||| ||  ||||||||||||||||
Qy        601 CGAGTGCGGGAGCTGCTGCTGGACGCCCTGAGCCCCGAGCAGCTAGTGCTCACCCTCCTG 660

Db       1108 gaagctgaaccacccaatgtgctggtgagccgtcccagcatgcccttcaccgaggcctcc 1167
              || ||||| ||| ||||||||| ||||| |||||    ||||||||||||||||||||
Qy        661 GAGGCTGAGCCGCCCCATGTGCTGATCAGCCGCCCCAGTGCGCCCTTCACCGAGGCCTCC 720

Db       1168 atgatgatgtccctcactaagctggcggacaaggaactggtgcacatgattggctgggcc 1227
              |||||||||||||| || ||| |||| ||||||||  ||| |||||||||  ||||||||
Qy        721 ATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTACACATGATCAGCTGGGCC 780

Db       1228 aagaaaatccctggctttgtggagctcagcctgttggaccaagtccggctcttagaaagc 1287
              ||||| || || ||||||||||||||||||||||| ||||||||| |||||||| || |||
Qy        781 AAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAAGTGCGGCTCTTGGAGAGC 840

Db       1288 tgctggatggaggtgctaatggtgggactgatgtggcgctcaatcgaccaccccggcaag 1347
              || ||||||||||||| ||||| |||| ||||||||||||||| || |||||||||||||||
Qy        841 TGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCAATTGACCACCCCGGCAAG 900

Db       1348 ctcattttcgctccagaccttgttctggacagggatgaggggaagtgcgtagaaggaatt 1407
              ||||| ||| |||||| || |||||||||||||||||||||| |||||||||||| |||
Qy        901 CTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGAAAATGCGTAGAAGGAATT 960

Db       1408 ctggaaatctttgacatgctcctggcgacgacgtcaaggttccgtgagttaaaactccag 1467
              |||||||||||||||||||||||||| || || |||||||| || |||||||||||||
Qy        961 CTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTTCGAGAGTTAAAACTCCAA 1020

Db       1468 cacaaggagtatctctgtgtgaaggccatgatcctgctcaactccagtatgtaccctttg 1527
              |||| || |||||||||| |||||||||||||||||| ||||| |||||||||||||  ||
Qy       1021 CACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAATTCCAGTATGTACCCTCTG 1080

Db       1528 gcttctgcaaaccaggaggccgaaagtagcaggaagctgacacacctactgaaccgggtg 1587
              |   | || | |||||| || || || |||||||||||| | |||  | ||||||| ||
Qy       1081 GTCACAGCGACCCAGGATGCTGACAGCAGCCGGAAGCTGGCTCACTTGCTGAACGCCGTG 1140

Db       1588 acagatgccctggtctgggtgattgcgaagagtggtatctcctcccagcagcagtcagtc 1647
              || |||||   |||| ||||||||||| ||||| || |||||||||||||||||| ||   |
Qy       1141 ACCGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATCTCCTCCCAGCAGCAATCCATG 1200

Db       1648 cgactggccaacctcctgatgctcctttctcacgtcaggcacatcagtaacaagggcatg 1707
              || ||||| ||||||||||||||||| || || ||||||||||      |||||||||||||
Qy       1201 CGCCTGGCTAACCTCCTGATGCTCCTGTCCCACGTCAGGCATGCGAGTAACAAGGGCATG 1260

Db       1708 gaacatctgctcagcatgaagtgcaaaaatgtggtcccggtgtatgacctgctgctggag 1767
              ||||||||||||| |||||||||||||||||||||||| |||||||||||||||||||||
Qy       1261 GAACATCTGCTCAACATGAAGTGCAAAAATGTGGTCCCAGTGTATGACCTGCTGCTGGAG 1320

Db       1768 atgctgaatgctcacacgcttcgaggctacaagtcctcaatctcggggtctgagtgcagc 1827
              ||||||||||| ||| |||||| |||| ||||||||| ||| |||||||| ||||||||
Qy       1321 ATGCTGAATGCCCACGTGCTTCGCGGGTGCAAGTCCTCCATCACGGGGTCCGAGTGCAGC 1380

Db       1828 tcaacagaggacagtaagaacaaagagagctcccagaacctacagtctcagtga 1881
               |  |||||||||||| | |||||||  ||||||||||||| |||||||||||||
Qy       1381 CCGGCAGAGGACAGTAAAAGCAAAGAGGGCTCCCAGAACCCACAGTCTCAGTGA 1434


RESULT    3
ID   T62844 standard; cDNA; 1458 BP.
AC   T62844;
DT   08-JUN-1997  (first entry)
DE   Mouse oestrogen receptor beta (ER-beta) cDNA.
KW   Orphan receptor; oestrogen receptor beta; ER-beta;
KW   nuclear receptor; prostate cancer; benign prostatic hyperplasia;
KW   ovary cancer; cardiovascular disease; osteoporosis;
KW   environment; pollutant; ss.
OS   Mus sp.
PN   WO9709348-A2.
PD   13-MAR-1997.
PF   09-SEP-1996; E03933.
PR   08-SEP-1995; GB-018272.
PR   15-MAR-1996; GB-005550.
PR   11-APR-1996; GB-007532.
PR   08-MAY-1996; GB-009[illegible].
PA   (KARO-) KARO BIO AB.
PI   Enmark E,  Gustafsson J,  Kuiper GG;
DR   WPI; 97-192842/17.
DR   P-PSDB; W14725.
PT   New isolated oestrogen receptor beta - used to develop prods. for
PT   treating e.g. cancers, CNS diseases, osteoporosis or cardiovascular
PT   disease
PS   Claim 6; Fig 14B; 45pp; English.
CC   A cDNA clone (T62844) codes for a novel mouse oestrogen receptor-
CC   related nuclear receptor (W13725), designated ER-beta.  Rat, human
CC   and mouse ER-beta (W14723-25) can be used to isolate molecules for
CC   use in the treatment of cardiovascular diseases, central nervous
CC   system diseases, osteoporosis, prostate or ovarian cancer or benign
CC   prostatic hyperplasia and to test environmental chemicals for
CC   oestrogenic activity.
SQ   Sequence   1458 BP;    365 A;    373 C;    403 G;    317 T;

  Query Match             69.2%;  Score 992;  DB 29;  Length 1458;
  Best Local Similarity   84.6%;  Pred. No. 0.00e+00;
  Matches [illegible];  Conservative    0;  Mismatches 221;  Indels    0;  Gaps    0;

Db         25 atgaactacagtgttcccagcagcaccggtaacctggaaggtgggcctgttccgcagact 84
              ||||| |||||  ||||||||     |  |||| |||||||||||||||| ||| ||||
Qy          1 ATGAATTACAGCATTCCCAGCAATGTCACTAACTTGGAAGGTGGGCCTGGTCGGCAGACC 60

Db         85 gcaagcccaaatgtgctatggccaacttctgggcacctgtctcctttagcgacccactgc 144
               |||||||||||||| | |||||||| |||| ||||| ||||||||||     || |  ||
Qy         61 ACAAGCCCAAATGTGTTGTGGCCAACACCTGGGCACCTTTCTCCTTTAGTGGTCCATCGC 120
```

```
Db    145 caatcatcgcttctctatgcagaacctcaaaagagtccttggtgtgaagcaagatcacta 204
          || | ||| | ||| ||||| ||||||||||||||||| ||||||||||||||||| |||
Qy    121 CAGTTATCACATCTGTATGCGGAACCTCAAAAGAGTCCCTGGTGTGAAGCAAGATCGCTA 180

Db    205 gaacacaccttgcctgtaaacagagagaccctgaagaggaagcttggcgggagcggttgt 264
          ||||||||||| ||||||||||||||||| ||||| |||||| || | |||| | ||||
Qy    181 GAACACACCTTACCTGTAAACAGAGAGACACTGAAAAGGAAGGTTAGTGGGAACCGTTGC 240

Db    265 gccagccctgttactagtccaagcaccaagagggatgctcacttctgtgccgtctgcagt 324
          ||||||||||||||| ||||| |  | ||||||||||||||||||||  | ||||||||
Qy    241 GCCAGCCCTGTTACTGGTCCAGGTTCAAAGAGGGATGCTCACTTCTGGCCTGTCTGCAGC 300

Db    325 gattatgcatctgggtatcattacggtgtctggtcctgtgaaggatgtaaggcctttttt 384
          ||||| ||||| || ||||| || || |||||||| ||||||||||||||||||||||||
Qy    301 GATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATGTAAGGCCTTTTTT 360

Db    385 aaaagaagcattcaaggacataatgactatatctgtccagccacgaatcagtgtacgata 444
          |||||||||||||||||||||||||| ||||| |||||||| || ||||||||||| ||
Qy    361 AAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCAGTGTACAATC 420

Db    445 gacaagaaccggcgtaaaaactgccaggcctgccgacttcgcaagtgttacgaagtagga 504
          || || |||||||| || | ||||||||||||||||||||| |||||||||||||| |||
Qy    421 GATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTACGAAGTGGGA 480

Db    505 atggtcaagtgtggatccaggagagaaaggtgtgggtaccgaatagtacgaagacagaga 564
          ||||| |||||||| ||| ||||||| || |||||||||||   | || || ||||||||
Qy    481 ATGGTGAAGTGTGGCTCCGGGAGACAGAGATGTGGGTACCGGCCTTGTGGGAGACAGAGA 540

Db    565 agtgccagcgagcaggtgcattgcctgaacaaagccaagagaaccagtgggcacacaccc 624
          ||||||  ||||||| |||| ||      ||| ||||||||||   | || |||    ||
Qy    541 AGTGCCGACGAGCAGCTGCACTGTGCCGGCAAGGCCAAGAGAAGTGGCGGCCACGGCGCC 600

Db    625 cgggtgaaggagctactgctgaactctctgagtcccgagcagctggtgctcaccctgctg 684
          || |||  |||||| |||||| || | ||||| ||||||||||| ||||||||||| |||
Qy    601 CGAGTGCGGGAGCTGCTGCTGGACGCCCTGAGCCCCGAGCAGCTAGTGCTCACCCTCCTG 660

Db    685 gaagctgatccacccaatgtgctagtgagtcgtcccagcatgcccttcaccgaggcctcc 744
          || |||||||| ||| |||||||    | || || |||||    ||||||||||||||||
Qy    661 GACGCTGACCCCCCCATGTGCTGATCAGCCCCCCAGTCCCGCCCTTCACCGAGGCCTCC 720

Db    745 atgatgatgtcccttacaaagctggctgacaaggaactgatccatatgattggctgggcc 804
          |||||||||||||| || ||| |||| ||||||||  || | ||||||||  ||||||||
Qy    721 ATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTACATATGATCAGCTGGGCC 780

Db    805 aagaaatccctggcctttgtggagctcagcctgttggaccaagtccgcctcttggaaagc 864
          |||||    |  | ||||||||||||||||||||| |||||||| || |||||||| |||
Qy    781 AAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAAGTGCGGCTCTTGGAGAGC 840

Db    865 tgctggatggaggtgctgatggtggggctaatgtggcgctccatcgaacaccccggcaag 924
          || |||||||||||| | ||| ||||||| ||||||||||| || || ||||||||||||
Qy    841 TGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCAATTGACCACCCCGGCAAG 900

Db    925 ctcatctttgctccagacctcgttctggacaggatgaggggaagtgcgtggaaagggatt 984
          ||||||||||||||||| || |||||||||||||||||||||| ||||| ||| || |||
Qy    901 CTCATCTTTGCTCCAGATCTTGTTCTGGACAGGATGAGGGGAAATGCGTAGAAGGGAATT 960

Db    985 ctggaaatctttgacatgctcctggcgacgacggcacggttccgtgagttaaaactgcag 1044
          |||||||||||||||||||||||||| || ||  || |||| || ||||||||||| ||
Qy    961 CTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTTCGAGAGTTAAAACTCCAA 1020

Db   1045 cacaaagaatatctgtgtgtgaaggccatgattctcctcaactccagtatgtaccacttg 1104
          |||||||||||||| ||||| ||||||||||| || ||||| |||||||||||||   ||
Qy   1021 CACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAATTCCAGTATGTACCCTCTG 1080

Db   1105 gctaccgcaagccaggaagcagaagatagccggaagctgaacacacttattgaagcagtg 1164
          |  ||  ||  |  |||||| || || || |||||||||| | ||| |   |||| |||
Qy   1081 GCTCACAGGACCCAGGATGCTGACAGCAGCGGAAGCTGGCTCACTTGCTGAACCGCCGTG 1140

Db   1165 acagatgccctggtctgggtgatttcaaaaagtaggatctcttcccagcagcagtcagtc 1224
          || |||||  |||| ||||||||| | |||||   | ||||| ||||||||||| ||  |
Qy   1141 ACCGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATCTCCTCCCAGCAGCAATCCATG 1200
```

```
Db   1225 cagtcagcgaacctcctgatgctcttctcatgtcagacacatcagttaacaaaaggcata 1284
          || |||||| |||||||||||||| || || || ||||||||     |||||||||||||
Qy   1201 CGCCTGGCTAACCTCCTGATGCTCCTGTCCCACGTCAGGCATGCCAGTAACAAGGGCATG 1260

Db   1285 gaacatctgctcagcatgaagtgcaaaaatgtggtcccggtgtacgacctgctgctggag 1344
          ||||||||||||| |||||||||||||||||||||||| ||||| |||||||||||||||
Qy   1261 GAACATCTGCTCAACATGAAGTGCAAAAATGTGGTCCCAGTGTATGACCTGCTGCTGGAG 1320

Db   1345 atgctgaatgctcacacgcttcgagggtacaagtcctcaatctcggggtctgggtgctgc 1404
          ||||||||||| |||  |||||| |||| ||||||||| ||| ||||||| | |||| ||
Qy   1321 ATGCTGAATGCCCACGTGCTTCGCGGGTGCAAGTCCTCCATCACGGGGTCCGAGTGCAGC 1380

Db   1405 tcgacagaggacagtaagagcaaagagggctcccagaacctccagtctcagtga 1458
           || ||||||||||||| ||||||||||||||||||||||  ||||||||||||
Qy   1381 CCGGCAGAGGACAGTAAAAGCAAAGAGGGCTCCCAGAACCCACAGTCTCAGTGA 1434
```

```
RESULT    4
ID   T06873 standard; DNA; 4963 BP.
AC   T06873;
DT   13-MAR-1996  (first entry)
DE   PhCMV*-1 promoter and human oestrogen receptor.
KW   Transactivator; tTA; tet operator; tetO; gene expression;
KW   tetracycline-responsive promoter; PhCMV*-1; human cytomegalovirus;
KW   transcription activation; eukaryotic cell; oestrogen receptor;
KW   ds; ss.
OS   Chimeric synthetic;
OS   Chimeric Homo sapiens.
PN   US5464758-A.
PD   07-NOV-1995.
PF   14-JUN-1993; 076726.
PR   14-JUN-1993; US-076726.
PA   (BUJA/) BUJARD H.
PA   (GOSS/) GOSSEN M.
PI   Bujard H,  Gossen M;
DR   WPI; 95-392012/50.
PT   Polynucleotide encoding trans-activator fusion protein contg. tet
PT   repressor - used to control expression of gene regulated by minimal
PT   promoter linked to tet operon, and vectors and cells where gene
PT   expression is regulated by tetracycline
PS   Disclosure; Fig 10A-B; 37pp; English.
CC   cDNA coding for the human oestrogen receptor was placed under
CC   control of the PhCMV*-1 promoter (see T06869), resulting in
CC   tetracycline-regulated expression vector pUHD10-3 contg. the
CC   sequence given in T06873.  The construct was transfected into
CC   eukaryotic cells expressing the tTA transactivator (T06867) to
CC   allow tetracycline-regulated expression of the receptor.
SQ   Sequence   4963 BP;   1259 A;   1300 C;   1278 G;   1126 T;

  Query Match             13.7%;  Score 197;  DB 17;  Length 4963;
  Best Local Similarity   66.7%;  Pred. No. 1.14e-121;
  Matches  424;  Conservative    0;  Mismatches  208;  Indels    4;  Gaps    2;

Db   1471 cccttcagtgaagcttggatgatgggcttactgaccaacctgcagacaaggaactgattt 1530
          |||||||  || || || ||||||    |  ||||||||  |||| |||| |||| ||||
Qy    703 CCCTTCACCGAGGCCTCCATGATGATGCCCCTGACCAAGCTGGCCGACAAGGACTTCCTA 762

Db   1531 cacatgatcaactgggcgaagagggtgccaggctttgtggatttgaccctccatgatcag 1590
          |||||||||| |||||| |||| | | || |||||||||||   | | |||     || ||
Qy    763 CACATGATCAACTGGGCCAAGAAGATCCCCGGCTTTGTGGAGCTCAGCCTCTTCCACCAA 822

Db   1591 gtccacttctagaatatgcctggctagagatcctgatgattggtctcgtctggcggctcc 1650
          || |  ||  | ||  |    ||| | ||| |   | ||||| || ||   | |||| ||
Qy    823 GTGCGGCTCTTGGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCA 882

Db   1651 atgaagcacccagtgaagctactgtttgctcctaacttgctcttggacaggaaccaggga 1710
          ||  | |||||    |||||    ||||||||  |  |  |  |||||||| |  |||| 
Qy    883 ATTGACCACCCCGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAGGCATGAGGGG 942

Db   1711 aaatgtgtagagggcatggtggagatcttcgacatgctgctggctacatcatctcggttc 1770
          ||||| ||||| || ||  |||| ||||| |||||||| ||||| ||  | ||  |||| 
Qy    943 AAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTT 1002
```

```
Db   1771 cgcatgatgaatctgcaggggagaggagtttgtgtgcctcaaatctattatttgcttaat 1830
Qy   1003 CGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAAT 1062

Db   1831 tctggagtgtacacatttctgtcccagaccctgaagtctctggaagagaaagaccatatc 1890
Qy   1063 TCCAGTATGTACCCTCTGGTCACAGGGACCCAGGATGCTGAC-AGCAGCCGGAAGCTGGC 1121

Db   1891 caccgagtcctggacaagatcacagacactttgatccacctgatggccaaggcaggcctg 1950
Qy   1122 TCACTTG--CTGAACGGCGTGACGGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATC 1179

Db   1951 accctgcagcagcagcaccagcggctggcccagctcctcctcatcctctcccacatcagg 2010
Qy   1180 TCCTCCCAGCAGCAATCCATGCGCCTGGCCAACCTCCTGATGCTCCTGTCCCACGTCAGG 1239

Db   2011 cacatgagtaacaaaggcatggagcatctgtacagcatgaagtgcaagaacgtggtgccc 2070
Qy   1240 CATGCGAGTAACAAGGGCATGGAACATCTGCTCAACATGAAGTGCAAAAATGTGGTCCCA 1299

Db   2071 ctctatgacctgctgctggagatgctggacgcccac 2106
Qy   1300 GTGTATGACCTGCTGCTGGAGATGCTGAATGCCCAC 1335

RESULT    5
ID   Q76270 standard; DNA; 4963 BP.
AC   Q76270;
DT   17-JUL-1995  (first entry)
DE   PhCMV*-1-controlled progesterone receptor
KW   tTA; transactivator; tetracycline-controllable transactivator;
KW   conditional inactivation; homologous recombination; gene expression;
KW   gene regulation; gene therapy; tetracycline-resistance; tetR;
KW   transgenic animal; PhCMV*-1; promoter; tetO; CMV;
KW   progesterone receptor; ds.
OS   Human cytomegalovirus Towne.
PN   WO9429442-A.
PD   22-DEC-1994.
PF   14-JUN-1994; U06734.
PR   14-JUN-1993; US-076327.
PA   (BADI ) BASF AG.
PI   Bujard H,  Gossen M,  Salfeld JG,  Voss JW;
DR   WPI; 95-036472/05.
PT   Regulatory systems using tetracycline-controllable transactivator
PT   (tTA) - useful for conditional inactivation or modulation of
PT   gene expression in a host cell or animal
PS   Disclosure; Page 54-58; 103pp; English.
CC   The minimal promoter PhCMV*-1 is based on the human CMV
CC   intermediate-early promoter (with enhancer region removed), fused
CC   to a tetracycline operator, tetO.  The promoter is used to
CC   express tTA transactivator in host cells.  In clone pUHD PGR4
CC   (given in Q76269) cDNA for rabbit progesterone receptor is under
CC   the control of PhCMV*-1.
SQ   Sequence   4963 BP;    1259 A;    1298 C;    1280 G;    1126 T;

  Query Match            13.7%;  Score 197;  DB 13;  Length 4963;
  Best Local Similarity  66.7%;  Pred. No. 1.14e-121;
  Matches  424;  Conservative    0;  Mismatches 209;  Indels   3;  Gaps   2;

Db   1471 cccttcagtgaagcttcgatgatgggcttactgaccaacctggcagacagggagctggtt 1530
Qy    703 CCCTTCACCGAGGCCTCCATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTA 762

Db   1531 cacatgatcaactgggcgaagagggtgccaggctttgtggatttgaccctccatgatcag 1590
Qy    763 CACATGATCAGCTGGGCCAAGAAGATCCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAA 822

Db   1591 gtccaccttctagaatgtacctggctaagatcctgatggattggtctcgtctagagctcc 1650
Qy    823 GTGCGGCTCCTCGAGAGCTGCTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCA 882

Db   1651 atggagcacccagtgaagctactgtttgctcctaacttgctcttggacaggaaccaggga 1710
Qy    883 ATTGACTACCTTTGCCAAGCTCATCTTCTCCCAGATCTTGTTCTGGACAGGCATGACGGG 942

Db   1711 aaatgtgtagagggcatggtggagatcttcgacatgctgctggctacatcatctcggttc 1770
Qy    943 AAATCCCTAGAAGGAATCTTCGAAATCTTTGACATGCTCCTGGCAACTACTTCTCGCTTT 1002

Db   1771 cgcatgatgaatctgcaggggagaggagtttgtgtgcctcaaatctattatttgcttaat 1830
Qy   1003 CGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAAT 1062

Db   1831 tctggagtgtacacatttctgtcccagaccctgaagtctctggaagagaaagaccatatc 1890
Qy   1063 TCCAGTATGTACCCTCTGGTCACAGGGACCCAGGATGCTGAC-AGCAGCCGGAAGCTGGC 1121

Db   1891 caccgagtcctggacaagatcacagacactttgatccacctgatggccaaggcaggcctg 1950
Qy   1122 TCACTTG--CTGAACGGCGTGACGGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATC 1179

Db   1951 accctgcagcagcagcaccagcggctggcccagctcctcctcatcctctcccacatcagg 2010
Qy   1180 TCCTCCCAGCAGCAATCCATGCGCCTGGCCAACCTCCTGATGCTCCTGTCCCACGTCAGG 1239

Db   2011 cacatgagtaacaaaggcatggagcatctgtacagcatgaagtgcaagaacgtggtgccc 2070
Qy   1240 CATGCGAGTAACAAGGGCATGGAACATCTGCTCAACATGAAGTGCAAAAATGTGGTCCCA 1299

Db   2071 ctctatgacctgctgctggagatgctggacgcccac 2105
Qy   1300 GTGTATGACCTGCTGCTGGAGATGCTGAATGCCCAC 1335

RESULT    6
ID   N70880 standard; cDNA; 2092 BP.
AC   N70880;
DT   24-FEB-1991  (first entry)
DE   cDNA encoding human oestrogen receptor protein.
KW   Human oestrogen receptor protein; ss cDNA; expression system; assay;
KW   steroid receptor protein;
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   CDS             293..1280
FT   /*tag=  a
PN   WO8705049-A.
PD   27-AUG-1987.
PF   18-FEB-1987; U00341.
PR   20-FEB-1986; US-833829.
PA   (CALB-) California Biotechnology Inc.
PI   Shine J;
DR   WPI; 87-250213/35.
DR   P-PSDB; P70543.
PT   Expression system for vertebrate steroid receptor protein - comprising
PT   DNA sequence encoding the protein linked to control sequences in
PT   eucaryotic hosts.
PS   Disclosure; Fig. 1-1 - 1-2; 26pp; English.
CC   The cDNA can be used in an expression system to express human
CC   oestrogen receptor protein.  The coding sequence is operably linked to
CC   control sequences compatible with eukaryotic host cells.  This method
CC   allows expression under conditions which favour appropriate post-
CC   translational processing.  It produces large amts. of purified protein
CC   useful in the design of agonist and antagonist cpds., for the study of
CC   the mechanism of action of the steroid binding proteins in general,
CC   and for use in diagnostic assays for the proteins or antibodies to
CC   them.  These assays are important in, eg the diagnosis of tumour
CC   sensitivities to steroid metabolism.  Suitable host cells are VERO,
CC   HeLa and CHO cells.
SQ   Sequence   2092 BP;    473 A;    607 C;    591 G;    421 T;

  Query Match            13.6%;  Score 195;  DB 2;  Length 2092;
  Best Local Similarity  66.5%;  Pred. No. 3.51e-120;
  Matches  421;  Conservative    0;  Mismatches 210;  Indels   3;  Gaps   2;

Db   1298 cccttcagtgaagcttcgatgatgggcttactgaccaacctggcagacagggagctggtt 1357
```

```
Qy     703 CCGTTCAGGGAGGCCTGCATCATGATGTCCCTGATCAATTTTTCCCACAAGGAGTTGGTA 762

Db    1358 cacatgatcaactgggcgaagagggtgccaggctttgtggatttgaccctccatgatcag 1417
           |||||||||| |||||| |||| | | || |||||||||||    | | |||     || ||
Qy     763 CACATGATCAGCTGGGCCAAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAA 822

Db    1418 gtccaccttctagaatgtgcctggctagagatcctgatgattggtctcgtctggcgctcc 1477
           || |  ||  | ||  |      ||| | ||| |   | ||||| || ||   | |||||||
Qy     823 GTGCGGCTCTTGGAGAGCTGTTCGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCA 882

Db    1478 atggagcacccagtgaagctactgtttgctcctaacttgctcttggacaggaaccaggga 1537
           || || ||||| |  |||||   | ||||||||   |   |   |||||||| |   ||||
Qy     883 ATTGACCACCCCGGCAAGCTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGG 942

Db    1538 aaatgtgtagagggcatggtggagatcttcgacatgctgctggctacatcatctcggttc 1597
           ||||| ||||| || ||  |||| ||||| |||||||| ||||| ||  | ||   ||||
Qy     943 AAATGCGTAGAAGGAATTCTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTT 1002

Db    1598 cgcatgatgaatctgcagggagaggagtttgtgtgcctcaaatctattatttgcttaat 1657
           ||    | | || || ||       | || | | | | ||  ||||   | || ||  |||| |||
Qy    1003 CGAGAGTTAAAACTCCAACACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAAT 1062

Db    1658 tctggagtgtacacatttctgtccagcaccctgaagtctctggaagagaaggaccatatc 1717
           ||  |  ||||| |  |  |  |     |||| | |  ||     | ||  |||     |  |
Qy    1063 TCCAGTATGTACCCTCTGGTCACAGCGACCCAGGATGCTGAC-AGCAGCCGGAAGCTGGC 1121

Db    1718 caccgagtcctggacaagatcacagacactttgatccacctgatggccaaggcacgcctg 1777
              |  |  ||| ||     | || ||  ||||| |        |||| ||||||    || |
Qy    1122 TCACTTG--CTGAACGCCGTGACCGATCCTTTGCTTTGCGTGATTCCCAAGAGCGGCATC 1179

Db    1778 accctgcagcagcagcaccagcggctggcccagctcctcctcatcctctcccacatcagg 1837
             ||    ||||||||     |    ||| |||||   | |||||    |  |||| |||||| |||||
Qy    1180 TCCTGCCAGCAGCAATCCATGCGGCTGGCTAACCTCCTGATGCTCCTGTCCCACGTCAGG 1239

Db    1838 cacatgagtaacaaaggcatggagcatctgtacagcatgaagtgcaagaacgtggtgccc 1897
           ||    ||||||||| |||||||| ||||||    || |||||||||||| || ||||| ||
Qy    1240 CATGCGAGTAACAAGGGCATGGAACATCTGCTCAACATGAAGTGCAAAAATGTGGTCCCA 1299

Db    1898 ctctatgacctgctgctggagatgctggacgcccac 1933
             | |||||||||||||||||||||||| | ||||||
Qy    1300 GTGTATGACCTGCTGCTGGAGATGCTGAATGCCCAC 1335


RESULT     7
ID   Q85976 standard; cDNA; 204 BP.
AC   Q85976;
DT   31-OCT-1995  (first entry)
DE   Human estrogen receptor cDNA.
KW   nuclear receptor protein; osteosarcoma; EAR-1r; generation; growth;
KW   differentiation; human; erb-A related gene; estrogen receptor; ds.
OS   Homo sapiens.
PN   J07033798-A.
PD   03-FEB-1995.
PF   20-JUL-1993; 200333.
PR   20-JUL-1993; JP-200333.
PA   (MITU ) MITSUBISHI KASEI CORP.
DR   WPI; 95-110650/15
DR   P-PSDB; R71566
PT   Nuclear receptor protein derived from osteo-blastic osteo-sarcoma
PT   - encoded by the EAR-1r gene, for investigation of generation of
PT   fat soluble active substances
PS   Disclosure; Page 6; 8pp; Japanese.
CC   This cDNA encodes an estrogen receptor (a member of a receptor
CC   superfamily). It is disclosed in the patent specification  The
CC   invention concerns an EAR-1r gene (human erb-A related gene)
CC   encoding a nuclear receptor protein (R71565) derived from human
CC   osteosarcoma (HOS) cells  The expression of EAR-1r is most abundant
CC   in the brain and was detectedd slightly in heart and placenta. The
CC   receptor protein may be used for the investigation of generation,
CC   differentiation and growth of fat soluble physiological substances
SQ   Sequence   204 BP;    59 A;    46 C;    55 G;    44 T;

  Query Match              8.9%;  Score 127;  DB 14,  Length 204;
  Best Local Similarity   81.3%;  Pred. No. 4.22e-70;
  Matches   165;  Conservative    0;  Mismatches   38;  Indels    0;  Gaps    0;

Db       1 tgtgcagtgtgcaatgactatgcttcaggctaccattatggagtctggtcctgtgaaggc 60
           || || || ||||  || || || || || || || |||||||||||||| ||||| ||
Qy     286 TGCGCTGTCTGCAGCGATTACGCATCGGATATCACTATGGAGTCTGGTCGTGTGAAGGA 345

Db      61 tgcaaggccttcttcaaaagaagtattcaaggacataacgactatatgtgtccagccacc 120
           || |||||||| || || ||||| ||||||||||||||| || ||||| |||||||| ||
Qy     346 TGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACA 405

Db     121 aaccagtgcaccattgataaaaacaggaggaagagctgccaggcctgccgactccgcaaa 180
           || ||||| || || ||||||||| || | |||||||||||||||||||| || || ||
Qy     406 AATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAG 465

Db     181 tgctacgaagtgggaatgatgaa 203
           || ||||||||||||||| ||||
Qy     466 TGTTACGAAGTGGGAATGGTGAA 488


RESULT     8
ID   Q10572 standard; DNA; 1047 BP.
AC   Q10572;
DT   09-APR-1991  (first entry)
DE   Human Natriuretic Peptide Receptor B.
KW   NPRB; ANP; BNP; CNP; kidney failure; heart failure; protein kinase;
KW   hyperaldosteronism; glaucoma; guanyl cyclase.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   Peptide         1..22
FT   /label= signal sequence
FT   Protein         12
FT   /label= mature NPBR
FT   Domain          23..455
FT   /label= extracellular domain
FT   /note= "binds natriuretic peptides A,B and C]"
FT   Domain          456..456
FT   /label= transmembrane domain
FT   Domain          479..1047
FT   /label= cytoplasmic domain
FT   /note= "GC and protien kinase activity"
FT   Modified -site  24  26
FT   /label= N-glycos_site
FT   Modified -site  35..37
FT   /label= N-glycos_site
FT   Modified -site  161..163
FT   /label= N-glycos_site
FT   Modified -site  195..197
FT   /label= N-glycos_site
FT   Modified -site  244..246
FT   /label= N-glycos_site
FT   Modified -site  277..279
FT   /label= N-glycos_site
FT   Modified -site  349..351
FT   /label= N-glycos_site
FT   Modified -site  600..602
FT   /label= N-glycos_site
PN   WO9100292-A.
PD   10-JAN-1991.
PF   22-JUN-1990; U03586.
PR   23-JUN-1989; US-370673.
PA   (GETH ) GENENTECH INC.
PI   Chang M, Goeddel D, Lowe D;
DR   WPI; 91-036711/05.
DR   N-PSDB; Q10324.
PT   Natriuretic protein receptor B - for diagnosis and treatment of
PT   kidney failure, heart failure, hyperaldosteronism, glaucoma etc.
PS   Claim 3; Fig 1; 49pp; English.
CC   The sequence was derived from the DNA encoding natriuretic peptide
CC   receptor B, NPRB, having guanyl cyclase (GC) activity and protein
CC   kinase activity.  The DNA can be inserted into expression vectors
CC   for the prodn  of the protein, opt. after being mutated to produce
CC   NPRB analogues. The protein has a mol wt. of 115 KD (calculated Mr-
CC   114,952).   The protein (or variants) can be used in treatment of
CC   natriuretic peptide disorders, and also to isolate peptides using
CC   affinity chromatography  Antibodies with affinity for NPRB can
CC   also be prepd.
SQ   Sequence   1047 BP;    87 A;    15 C;    83 G;    51 T;

  Query Match              6.3%;  Score 91;  DB 2;  Length 1047,
  Best Local Similarity   8.9%;  Pred. No. 2.62e-44;
  Matches   83;  Conservative  269;  Mismatches 572;  Indels  10;  Gaps  10;

Db     108 varnashwrnnnntagavasgnsakndhyrtnvrtgnsankngnnvvtnhghnnwtaraa 167
            :|: |:: :       ||:| |  |: .. .| ..    .  :    :: : : :
Qy     198 AAACAGAGAGATACTGAAAAGGAAGGTTAGTGGGAATTGTTGGGCAGCCCTGTTACTGG 257

Db     168 nnyndart-ddrnhyntnrgvnnanngsnnsvnhnvyarnnggnnnathnnrangrnvyn 226
              : :  | :.. .      |.  |  .  .. .  . . |       .  . |. ..
Qy     258 TCCAGGTTCAAAGAGGGATGCTCACTTCTGCGCTGTCTGCAGCGATTACGCATCGGGATA 317

Db     227 cgnnnmnhnnnnnanrnnntn-gdyvnnyndvngnsnragntratgrnwndnrtrnnana 285
                :          :    |  |: :  : :: | :       | |  : : : :|    |
Qy     318 TCACTATGGAGTCTGGTCGTGTGAAGGATGTAAGGCCTTTTTAAAAGAAGCATTCAAGG 377

Db     286 nrnanntvnvntyrnnnnnnynnnnnrnnnrarndrgvnngnsnmnnnagnnydgnnnya 345
                  |:   ::      :              : :  :     : :   |   ::|    |
Qy     378 ACATAATGATTATATTTGTCCAGCTACAAATCAGTGTACAATCGATAAAAACCGGCGCAA 437

Db     346 nvnnntnnnggtrndgnrnvnkmngrryhgvtgnvvmdknndrntdnvnwamgdndsgdn 405
             :        ||    :   :  : :: |:: :| |  ::...   ..  .   .|  . .. .
Qy     438 GAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTACGAAGTCGGAATGGTGAAGTCTGGCTC 497

Db     406 nnsahysgankrnwwtgrnnnwvkganssdnnncanlnddnscdktnnstnanvangtgn 465
                :  ||     : ||:     ::      .     |          |      |   |   |
Qy     498 CCGGAGAGAGAGATGTGGGTACCGCCTTGTGCGGAGACAGAGAAGTGCC-GACGAGCAGC 556

Db     466 tnnmngvssnnnnrknmnnknnasmnwrnrwnnnnngnsnryhkgagsrntnsnrgssyg 525
           |  :  :  :    :   :      |:: :: ::           ::      :: | : :| ::
Qy     557 TGCACTGTGCCG-GCAAGGCCAAGAGAA-GTGGCCGCTACGCGCCCGAGTCCGGGAGCT 614

Db     526 snmtahgkynnnantghnkgnvvankhvnkkrnnntrnvnnnnkhmrdvnnnhntrnnga 585
           :     ::|:         ||:      ::|  :: :  :   |  :        : :::
Qy     615 GCTGCTGGACGCCCTGAGCCCCCAGCAGCTAGTGCTCACCCTCCTGGAGGCTGAGCCGCC 674

Db     586 gndnnnngnvtnycnrgsnndnnnndsnnndwmnrysnnndnvkgmannhnsnnsshgsn 645
           : :       | :  :      :       :.     .  ...    . .. .|  : :  : : :
Qy     675 CCATGTGCTGATCAGCCGCCCCAGTGCGCCTTTCACCGAGGCCTCCATGATGATGTCCCT 734

Db     646 kssncvvdsrnvnkntdygnasn-rstannddnnanyakknntannnnsgnnnnttgmna 704
           : :  ::: : : :  : | |:  ::||    .  | :|.             :      |  :
Qy     735 GACCAAGTTGGCCGACAAGGAGTTGGTACACATGATCAGCTGGGCCAAGAAGATTCCCGG 794

Db     705 advysngnnnnnnanrsgnnynngndnsnknnvnkvrngnrnynrnsndrtnnnnnnvnn 764
             : :: |         | ::    :  |      :  : ::    :   : : : |       :
Qy     795 CTTTGTGGAGCTCA-GCCTGTTCGACCAAGTGCGGCTCTTGGAGAGCTGTTGGATGGAGG 853

Db     765 mnrcwandnanrndngnnkgnnrrnnkncgtsnndnnnnrmnnyannnnknvnnrtnayn 824
                :| :   : :    :|  :   :    |         :  :|            |  :
Qy     854 TGTTAATGATGGGGCTGATGTGGCGCTCAATTGACCACCCCGGCAAGCTCATCTTTGCTC 913

Db     825 nnkrkananmynnnnhsvannnkrgntvnanandsvtnynsdnvgntans-anstnmnvv 883
             :::     :    :::     ::    : | | :::|     : :  |  : |  |      :
Qy     914 CAGATCTTGTTCTGGACAGGGATGAGGGGAAATGCGTAGAAGGAATTCTGGAAATCTTTG 973

Db     884 tnnndnytcndanndnndvykvntngdaymvvsgnngrngnrhannnarmanandavss 943
                : : | :    :  ::::.     |. .....     .    :      ..| |  . . .
Qy     974 ACATGCTCCTGGCAACTACTTCAAGGTTTCGAGAGTTAAAACTCCAACACAAAGAATATC 1033

Db     944 nrnrhrnhdnnrn-rngvhtgnvcagvvgnkmnrycnngdtvntasrmnsngnankhvs 1002
              ::: ..   .      |      |  . .   .             |          .
Qy    1034 TCTGTGTCAAGGCCATGATCCTGCTCAATTCCAGTATGTACCCTCTGGTCACAGCGACCC 1093

Db    1003 -sttkdandnngcnnnnnrgdvnmkgkgkmrtyw 1035
             :  ::  :          :|:: ::|: ::: ::
Qy    1094 AGGATGCTGACAGCAGCCGGAAGCTGGCTCACTT 1127
```

```
RESULT    9
ID   Q10572 standard; DNA; 1047 BP.
AC   Q10572;
DT   09-APR-1991  (first entry)
DE   Human Natriuretic Peptide Receptor B.
KW   NPRB, ANF, BNF, CNF, kidney failure, heart failure, protein kinase;
KW   hyperaldosteronism, glaucoma, guanyl cyclase.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   Peptide         1..22
FT   /label= signal sequence
FT   Protein         12
FT   /label= mature NPBR
FT   Domain          23..455
FT   /label= extracellular domain
FT   /note= "binds natriuretic peptides A,B and C]"
FT   Domain          456..456
FT   /label= transmembrane domain
FT   Domain          479..1047
FT   /label= cytoplasmic domain
FT   /note= "GC and protien kinase activity"
FT   Modified -site  24..26
FT   /label= N-glycos_site
FT   Modified -site  35..37
FT   /label= N-glycos_site
FT   Modified -site  161..163
FT   /label= N-glycos_site
FT   Modified -site  195..197
FT   /label- N-glycos_site
FT   Modified -site  244..246
FT   /label= N-glycos_site
FT   Modified -site  277..279
FT   /label- N-glycos_site
FT   Modified -site  349..351
FT   /label- N-glycos_site
FT   Modified -site  600..602
FT   /label- N-glycos_site
PN   WO9100292-A.
PD   10-JAN-1991.
PF   22-JUN-1990; U03586.
PR   23-JUN-1989; US-370673.
PA   (GETH ) GENENTECH INC.
PI   Chang M, Goeddel D, Lowe D;
DR   WPI; 91-036711/05.
DR   N-PSDB, Q10324.
PT   Natriuretic protein receptor B - for diagnosis and treatment of
PT   kidney failure, heart failure, hyperaldosteronism, glaucoma etc.
PS   Claim 3; Fig 1; 49pp; English.
CC   The sequence was derived from the DNA encoding natriuretic peptide
CC   receptor B, NPPB, having guanyl cyclase (GC) activity and protein
CC   kinase activity  The DNA can be inserted into expression vectors
CC   for the prodn. of the protein, opt. after being mutated to produce
CC   NPPB analogues  The protein has a mol wt. of 120 kD (calculated Mr=
CC   114,952).  The protein (or variants) can be used in treatment of
CC   natriuretic peptide disorders, and also to isolate peptides using
CC   affinity chromatography.  Antibodies with affinity for NPRB can
CC   also be prepd.
SQ   Sequence   1047 BP;    87 A;    15 C;    83 G;    51 T;

  Query Match              5.0%;  Score 71,  DB 2,  Length 1047;
  Best Local Similarity   9.3%;  Pred. No. 2.10e-30;
  Matches   64;  Conservative  203;  Mismatches 415;  Indels   9;  Gaps   9;

Db     340 gnnnyanvnnntnnnggtrndgnrnvnkmngrryhgvtgnvvmdknndrntdnvnwamgd 399
             |   .        |  ||    |                   ||         :  |  : : : :
Qy    1371 GGACCCCGTGATGGAGGACTTGCACCCCGAAGCACGTGGGCATTCAGCATCTCCACCAG 1312

Db     400 ndsgdnnna-shysg-ankrnwwtgrnnnwvkgannsdnnn-canlnddnscdktnnstn 456
              ::|:    | |...| |    ..|      ...     .:     || : :: :':: :.
Qy    1311 CAGGTCATACACTGGGACCACATTTTTGCACTTCATGTTGAGCAGATGTTCCATGCCCTT 1252
```

```
Db      457 anrnnqtqntnnmnqvssnnnrknmnnknnasmnwrnrwnnnnngnsnryhkgngsrnt 516
               |   |    :  : :    :: :  :   ::  . ..     | . .    |    |
Cp     1251 GTTACTCGCATGCCTGACGTG-GGACAGGAG-CATCAGGAGGTTAGCCAGGCGCATGGAT 1194

Db      517 nanrqssyqnntnkqkynnnntqhnkqnvvnnkhvnkkrnnntrnvnnnkhnrdvnn 576
             .  |. .|. .  :| .     | .. | ..   ..   .    : .    .... :
Cp     1193 TGCTGC-TGGGAGGAGATGCCGCTCTTGGCAATCACCCAAACCAAAGCATCGGTCACGGC 1135

Db      577 nhntrnnqacndnnnncvvtryvnrqsnndnnnrdsnnndwmnrysnnndnvkgmannhn 636
             :  :   |  :      :| :| :    :    ::    .  ...   .  . .|
Cp     1134 GTTCAGCAAGTGAGCCAGCTTCCGGCTGCTGTCAGCATCCTGGGTCGCTGTGACCAGAGG 1075

Db      637 snnsshqsnkssnvvdsrnvnkntdygnasnrstannddnnanyakknntannnnsgnn 696
            :  : : : ::    :::  :    ::|  :  :    ::  | :|::  |      |
Cp     1074 GTACATACTGGAATTGAGCAGGATCATGGCCTTGACACAGAGA-TATTCTTTGTGTTGGA 1016

Db      697 nnttqmnaadvysngnnnnnanrsgnnynngndnsnknnvnkvrngnrnynrnsndrtn 756
              || :     ::        | :    :    :      :  :: |   :   : : |
Cp     1015 GTTTTAACTCTCGAAACCTTGAAGTAGTTGCCAGGAGCATGTCAAAGATTTCCAGAATTC 956

Db      757 nnnnnvnnmnrcwandnanrndnqnnkgnnrrnnknggtsnndnnnnrmnnyannnnknv 816
                 :  :   :      :      :   ::    | |:  :    ::  :     : :
Cp      955 CTTCTACGCATTTCCCCTCATCCCTGTCCAGAACAAGATCTGGAGCAAACATGAGCTTGC 896

Db      817 nnrtnnaynnnkrknnnnnynnnnhsvannnkrgntvnanandsvtnynsdnvgntansan 876
              :        ::  |       : :|        :   | :::| :    :
Cp      895 CGGGGTGGTCAATTGAGCCCCACATCAGCCCCATCATTAACACCTCCATCCAACAGCTCT 836

Db      877 stnmnvvtnnndnytcndnnndnndvykvntngdaymvvsgnngrngnrhannnarmana 936
            :  : ::    :: |  :   :  :::::       ::: |  |:    :     :
Cp      835 CCAAGAGCCGCACTTGGTCGAACAGGCTGAGCTCCACAAAGCCGGGAATCTTCTTGGCCC 776

Db      937 nndavssnrnrhrnhdnnrnrnqvhtgnvcaqvvqnkmnrycnngdtvntasrmnsngna 996
                :   : : :: :: :    ::|| :   ::  :   ::|  |: : | ::   | |
Cp      775 AGCTCATCATGTG-TACCAACTCCTTGTCGGCCAACTTGCTCAGGACAT-CATCATGGA 718

Db      997 nknhvssttkdandnnqcnnnnnrqdvnmkq 1027
             : : :: |::| :  ||     :| : ::|
Cp      717 GGCCTCCGTGAAGGGCGCACTGGGCGGCTG 687


RESULT   10
ID   T42193 standard; DNA; 3460 BP.
AC   T42193;
DT   20-JAN-1997  (first entry)
DE   Fragment of pig oestrogen receptor gene
KW   Pig; oestrogen receptor; exon; polymorphism; allele; RFLP; litter size;
KW   restriction fragment length polymorphism; Chinese; American; PCR; primer;
KW   amplification; polymerase chain reaction; ds.
OS   Sus scrofa.
FH   Key             Location/Qualifiers
FT   misc_feature    303..304
FT   /*tag=  a
FT   /note= "undetermined length gap missing in DNA sequence
FT   between these nucleotides"
FT   exon            534..600
FT   /*tag=  b
FT   /note= "exon 3"
FT   misc_feature    1197..1198
FT   /*tag=  c
FT   /note= "approx  100-200 bases missing from DNA sequence
FT   between these nucleotides"
FT   primer_bind     1898..1915
FT   /*tag=  d
FT   /note= "binds primer A1CFII (T42200)"
FT   primer_bind     2011..2031
FT   /*tag=  e
FT   /note= "binds primer DVO28 (T42198)"
FT   primer_bind     2146..2171
FT   /*tag=  f
FT   /note= "binds primer ESR REV (T42195)"
FT   primer_bind     2636
FT   /*tag=  g
```

```
FT   primer_bind     2636..2663
FT   /*tag=  h
FT   /note= "binds primer ESRSR (T42197)"
FT   variation       2688..2693
FT   /*tag=  i
FT   /note= "polymorphic PvuII site"
FT   primer_bind     2729..2755
FT   /*tag=  j
FT   /note= "binds primer ESRSF (T42196)"
FT   primer_bind     2922..2940
FT   /*tag=  k
FT   /note= "binds primer B1F (T42201)"
FT   primer_bind     2982..3002
FT   /*tag=  l
FT   /note= "binds primer DVO29 (T42199)"
FT   primer_bind     3331..3355
FT   /*tag=  m
FT   /note= "binds primer ESR FOR (T42194)"
FT   misc_feature    3123..3128
FT   /*tag=  n
FT   /note= "constant PvuII site"
PN   US5550024-A.
PD   27-AUG-1996.
PF   19-APR-1991; 687708.
PR   19-APR-1991; US-687708.
PR   19-OCT-1992; US-961819.
PR   26-SEP-1994; US-312312.
PA   (BIOT-) BIOTECHNOLOGY RES & DEV CORP.
PA   (IOWA ) UNIV IOWA STATE RES FOUND INC.
PI   Jacobson CD,  Mileham AJ,  Plastow GS,  Rothschild MF;
PI   Tuggle CK,  Vaske DA;
DR   WPI; 96-401592/40.
PT   Screening pigs for high litter size - by identifying a polymorphism
PT   in the pig oestrogen receptor gene in a DNA sample
PS   Disclosure; Column 19-24; 26pp; English.
CC   This is a fragment of the pig oestrogen receptor (ESR) gene including
CC   exon 3.  The sequence covers the region contg. a polymorphic PvuII
CC   restriction enzyme site which is used as site for determining the ESR
CC   allele carried by the animal.  Restriction fragment length polymorphism
CC   (RFLP) analysis of this region revealed two alleles of 4.3 and 3.7 kb.
CC   Both alleles are involved with increased pig litter size dependent on the
CC   breed of pig involved e.g. Chinese or American pigs.  However pigs with
CC   the 3.7 kb allele had between 1.78-2.4 more pigs born in each litter and
CC   1.8-2.4 more pigs born alive than pigs with the 4.3 kb allele. The allele
CC   can also be determined by PCR using the primers (T42194-T42201).
SQ   Sequence   3460 BP;    1052 A;    765 C;    634 G;    1009 T;

  Query Match             5.0%;  Score 71;  DB 24;  Length 3460;
  Best Local Similarity  83.3%;  Pred. No. 2.10e-30;
  Matches    95,  Conservative    0,  Mismatches  18;  Indels   1;  Gaps    1;

Db      490 ttcatcatgcccacttcgtagcacttgcgtagcccggcaggc tggcagctcttcctcctg 548
            |||| ||| |||||||||| ||||| || || |||||||| || |||||||| | || |
Cp      488 TTCACCATTCCCACTTCGTAACACTTCCGAAGTCGGCAGGCCTGGCAGCTCTTGCGCCGG 429

Db      549 ttcttatcaattgtgcactggttggtagctggacacatgtaatcattatgtcct 602
            || ||||| ||||| ||||| || ||||||||||| || || ||||||||||||
Cp      428 TTTTTATCGATTGTACACTGATTTGTAGCTGGACAAATATAATCATTATGTCCT 375


RESULT   11
ID   N80922 standard; cDNA; 2153 BP.
AC   N80922;
DT   29-DEC-1990  (first entry)
DE   Sequence encoding human estrogen related receptor protein (hERR2)
KW   Hormone receptor; hormone-binding; transcription activation; ss.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   CDS             100..1401
FT   /*tag=  a
PN   WO8803168-A.
PD   05-MAY-1988.
PF   23-OCT-1987; [illegible]02782.
```

```
PR   4-OCT-1986; US-922585.
PR   20-OCT-1987; US-108471
PA   (SALK) Salk Inst for Biol Stud.
PI   Evans RM, Weinberger CA  Hollenberg SM, Giguere V
DR   WPI; 88-133242/19.
DR   P-PSDB; P80931.
PT   Recombinant DNA encoding hormone receptors -
PT   comprising glucocorticoid  mineralocorticoid  thyroid hormone
PT   and novel hormone receptors
PS   Claim 16; Fig V-2(B)-1 and -2; 243pp; English.
CC   DNAs encoding hormone receptors and the hormone receptors themselves are
CC   claimed. The DNA can be used to make the hormone receptor proteins and
CC   functional modified forms in quantities not previously possible. The
CC   receptor proteins can be used to screen cpds. for receptor-agonist or
CC   receptor-antagonist activity. They can also be used in diagnostic assays.
CC   Also claimed is pure DNA encoding oestrogen-related receptors hERR1 and
CC   hERR2 and hERR1 and hERR2  The new hERR1 and hERR2 receptors will provide
CC   the basis for development of an assay system that will lead to the
CC   identification of novel hormones.
SQ   Sequence   2153 BP;   445 A;   627 C;   624 G,   457 T,

  Query Match              4.7%;  Score 68;  DB 2;  Length 2153,
  Best Local Similarity  67.2%;  Pred. No. 2.33e-28;
  Matches   133;  Conservative    0;  Mismatches  65;  Indels   0,  Gaps   0,

Db     412 gtgtgcggggacattgcttctggctacgactatggagtggtctgtgaaggttgcaag 471
            || ||| | ||    || || || || ||||||||||    || || || |  || |||
Qy     292 GTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATCTAAG 351

Db     472 gcgttcttcaagagaaccattcaaggaaacatcgaatacagctgccctgccaccaacgag 531
            || || || || |||| ||||||||| | |  || || |  || || || || ||  ||
Qy     352 GCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCAG 411

Db     532 tgtgaggtcgacaagatcaggaggaagtcctgtcaggcctgccggttcatgaaatgcctc 591
            | |   ||   ||    |||||||  ||| ||||||||||  |   ||| ||    |
Qy     412 TGTACAATCGATAAAAACCGGCGCAAGAGTTGCCAGGCCTGCCGACTTCGGAAGTGTTAC 471

Db     592 aaagtggggatgctgaag 609
               |||||| ||| |||||
Qy     472 GAAGTGGGAATGGTGAAG 489


RESULT   12
ID   Q76269 standard; DNA; 6244 BP
AC   Q76269;
DT   17-JUL-1995  (first entry)
DE   PhCMV*-1-controlled progesterone receptor
KW   tTA; transactivator; tetracycline-controllable transactivator,
KW   conditional inactivation  homologous recombination, gene expression,
KW   gene regulation; gene therapy; tetracycline-resistance; tetR,
KW   transgenic animal  PhCMV*-1, promoter, tetO  CMV,
KW   progesterone receptor; ds.
OS   Human cytomegalovirus Towne.
PN   WO9429442-A
PD   22-DEC-1994.
PF   14-JUN-1994; U06734.
PR   14-JUN-1993; US-076327
PA   (BADI ) BASF AG.
PI   Bujard H   Gossen M   Salfeld JG,  Voss JW;
DR   WPI; 95-036472/05.
PT   Regulatory systems using tetracycline-controllable transactivator
PT   (tTA)  - useful for conditional inactivation or modulation of
PT   gene expression in a host cell or animal
PS   Disclosure; Page 54-58; 103pp; English.
CC   The minimal promoter PhCMV*-1 is based on the human CMV
CC   intermediate-early promoter (with enhancer region removed), fused
CC   to a tetracycline operator, tetO.  The promoter is used to
CC   express tTA transactivator in host cells.  In clone pUHD BGR3
CC   (given in Q76269), cDNA for rabbit progesterone receptor is under
CC   the control of PhCMV*-1.
SQ   Sequence   6244 BP;   1483 A;   1749 C;   1610 G,   1402 T,

  Query Match              4.4%;  Score 63;  DB 13;  Length 6244;
  Best Local Similarity  66.3%;  Pred. No. 5.58e-25;
  Matches   128,  Conservative    0,  Mismatches  65,  Indels   0,  Gaps   0;

Db    2365 tctgtggggatgaagcatcaggctgtcattatggtgtcctcacctgtgggagctgtaagg 2424
            ||||  | ||| | ||||| || | ||| ||||| |||     | ||||    | |||||||
Qy     293 TCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATCTAAGG 352

Db    2425 tcttctttaaaaggcaatggaagggcagcataactatttatgtgctaaaagaaatgact 2484
             ||| ||||||||     ||   |||| ||   || | ||| | ||| | |   | |||| | |
Qy     353 CCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCAGT 412

Db    2485 gcattgttgataaaatcagcaggaaaactgccggcgtgtcgacttagaaaatgctgtc 2544
            | |   | ||||||| ||   | || | ||||| ||| || || ||| | ||||| |
Qy     413 GTACAATCGATAAAAACCGGCGCAAGAGTTGCCAGGCCTGCCGACTTCGGAAGTGTTACG 472

Db    2545 aagctggcatggt 2557
            |||  || |||||
Qy     473 AAGTGGGAATGGT 485


RESULT   13
ID   T06872 standard, DNA, 6244 BP.
AC   T06872;
DT   13-MAR-1996  (first entry)
DE   PhCMV*-1 promoter and rabbit progesterone receptor.
KW   Transactivator, tTA, tet operator, tetO, gene expression;
KW   tetracycline-responsive promoter, PhCMV*-1, human cytomegalovirus;
KW   transcription activation, eukaryotic cell, progesterone receptor;
KW   ds; ss.
OS   Chimeric synthetic;
OS   Chimeric Oryctolagus cuniculus.
PN   US5464758-A.
PD   07-NOV-1995.
PF   14-JUN-1993, 076726.
PR   14-JUN-1993; US-076726.
PA   (BUJA/) BUJARD H.
PA   (GOSS/) GOSSEN M.
PI   Bujard H,  Gossen M;
DR   WPI; 95-392612/50.
PT   Polynucleotide encoding trans activator fusion protein contg. tet
PT   repressor  - used to control expression of gene regulated by minimal
PT   promoter linked to tet operon, and vectors and cells where gene
PT   expression is regulated by tetracycline
PS   Disclosure, Fig 9A-C, 37pp, English.
CC   cDNA coding for the rabbit progesterone receptor was placed under
CC   control of the PhCMV*-1 promoter (see T06869), resulting in
CC   tetracycline-regulated expression vector pUHDpgr-3 contg  the
CC   sequence given in T06872.  The construct was transfected into
CC   eukaryotic cells expressing the tTA transactivator (T06867) to
CC   allow tetracycline-regulated expression of the receptor.
SQ   Sequence   6244 BP,   1483 A,   1750 C,   1609 G,   1402 T;

  Query Match              4.4%;  Score 63;  DB 17;  Length 6244;
  Best Local Similarity  66.3%;  Pred. No. 5.58e-25;
  Matches   128;  Conservative    0,  Mismatches  65,  Indels   0,  Gaps   0;

Db    2365 tctgtggggatgaagcatcaggctgtcattatggtgtcctcacctgtgggagctgtaagg 2424
            ||||  | ||| | ||||| || | ||| ||||| |||     | ||||    | |||  |||
Qy     293 TCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATCTAAGG 352

Db    2425 tcttctttaaaaggcaatggaagggcagcataactatttatgtgctaaaagaaatgact 2484
             ||| ||||||||     ||   |||| ||   || | ||| | ||| | |   | |||  | |
Qy     353 CCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCAGT 412

Db    2485 gcattgttgataaaatcagcaggaaaactgccggcgtgtcgacttagaaaatgctgtc 2544
            | |   | ||||||| ||   | || | ||||| ||| || || ||      ||||
Qy     413 GTACAATCGATAAAAACCGGCGCAAGAGTTGCCAGGCCTGCCGACTTCGGAAGTGTTACG 472

Db    2545 aagctggcatggt 2557
            |||  || |||||
Qy     473 AAGTGGGAATGGT 485
```

```
RESULT   14
ID   Q79934 standard; cDNA; 2992 BP.
AC   Q79934;
DT   01-AUG-1995  (first entry)
DE   Human liver hap cDNA.
KW   RAR-beta; retinoic acid receptor beta; hap gene; hepatoma;
KW   hepatocellular carcinoma; hepatocyte; liver; hepatitis B virus,
KW   ss.
OS   Homo sapiens.
FH   Key                 Location/Qualifiers
FT   misc_signal         119..121
FT   /*tag=  a
FT   /note= "untranslated start codon"
FT   misc_signal         126..128
FT   /*tag=  b
FT   /note= "untranslated start codon"
FT   misc_signal         178..180
FT   /*tag=  c
FT   /note= "untranslated start codon"
FT   misc_signal         237..239
FT   /*tag=  d
FT   /note= "untranslated start codon"
FT   CDS                 322..1668
FT   /*tag=  e
FT   /product= RAR-beta
FT   polyA_signal        1954..2959
FT   /*tag=  f
PN   US5376530-A.
PD   27-DEC-1994.
PF   16-DEC-1987; 133687
PR   16-DEC-1987; US-133687.
PR   17-DEC-1987; US-134130.
PR   20-JUN-1988; US-209009.
PR   30-NOV-1988; US-278136.
PR   30-MAR-1989; US-330405.
PR   21-AUG-1991; US-751612.
PR   30-MAR-1992; US-860577
PR   11-DEC-1992; US-989902.
PR   22-JUL-1993; US-095706.
PA   (INSP ) INST PASTEUR.
PI   Brand N,  Chambon P,  De The HB,  Dejean A,  Krust A;
PI   Marchio A, Petkovich M,  Tiollais P;
DR   WPI; 95-043461/06
DR   P-PSDB; R66764.
PT   Detection of transformed hepatocyte(s) - by detecting a 2.5 kb
PT   mRNA species in hepatic tissue using DNA encoding a retinoic acid
PT   receptor-beta fragment.
PS   Disclosure; Fig 2A 2B; 35pp; English
CC   A hepatitis B virus integration in a DNA fragment linked to
CC   hepatocellular carcinoma was used as a probe to clone cDNA from a
CC   human liver cDNA library.  The cellular gene, named hap, encoded a
CC   retinoic acid receptor identified as RAR-beta
SQ   Sequence   2992 BP;    899 A;    652 C;    600 G;    841 T;

  Query Match             4.1%;  Score 59;  DB 13;  Length 2992;
  Best Local Similarity  64.8%;  Pred. No. 2.63e-22;
  Matches  129;  Conservative    0;  Mismatches  70;  Indels   0;  Gaps   0;

Db    561 ctgcttcgtctgccaggacaaatcatcagggtaccactatggggtcagcgcctgtgaggg 620
          ||||   ||||||   ||  |  |||| || || |||||||| ||| |  | ||||| ||
Qy    285 CTGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGG 344

Db    621 atgtaagggctttttccgcagaagtattcagaagaatatgatttacacttgtcacccaga 680
          |||||||| ||||||    ||||| |||||     |||    ||| | |||||   |
Qy    345 ATGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTAC 404

Db    681 taagaactgtgttattaataaagtcaccaggaatcgatgccaatactgtcgactccagaa 740
           ||  | |||   ||  |||||  |    | ||  | |||||   ||| ||||| | |||
Qy    405 AAATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAA 464

Db    741 gtgctttgaagtgggaatg 759
          ||| |  ||||||||||||
Qy    465 GTGTTACGAAGTGGGAATG 483
```

```
RESULT   15
ID   Q65572 standard; cDNA; 2992 BP
AC   Q65572;
DT   11-JAN-1995  (first entry)
DE   Human liver hap cDNA.
KW   Liver; hap; retinoic acid receptor; steroid; thyroid; hormone;
KW   hepatoma; retinoid; antibody; ss.
OS   Homo sapiens.
FH   Key                 Location/Qualifiers
FT   CDS                 322..1668
FT   /*tag=  a
FT   /product= RAR-beta retinoic acid receptor.
FT   misc_difference 916..918
FT   /*tag=  b
FT   /transl_except= GCT encodes Glycine.
FT   misc_difference 1372..1374
FT   /*tag=  c
FT   /transl_except= GGA encodes Alanine.
PN   US5317090-A.
PD   31-MAY-1994.
PF   16-DEC-1987; 133687.
PR   16-DEC-1987; US-133687.
PR   17-DEC-1987; US-134130.
PR   20-JUN-1988; US-209009.
PR   30-NOV-1988; US-278136.
PR   30-MAR-1989; US-330405.
PR   21-AUG-1991; US-751612.
PR   30-MAR-1992; US-860577.
PR   11-DEC-1992; US-989902.
PA   (INSP ) INST PASTEUR.
PI   Brand N,  Chambon P,  De THE HB,  Dejean A,  Krust A;
PI   Marchio A, Petkovich M,  Tiollais P;
DR   WPI; 94-176333/21.
PT   Antibody specific for retinoic acid receptor beta - useful for
PT   detecting, quantifying and identifying agonists and antagonists
PT   of retinoid activity.
PS   Disclosure; Figure 2; 35pp; English.
CC   The hap cDNA encodes a retinoic acid receptor called RAR-beta.  The
CC   hap gene is transcribed at low level in most human tissues, but the
CC   gene is overexpressed in prosate and kidney.  Six out of seven
CC   hepatoma or hepatoma-derived cell lines express a small hap
CC   transcript which is undetectable in normal adult and foetal livers
CC   but present in all non-hepatic tissues tested.
SQ   Sequence   2992 BP;    899 A;    653 C;    599 G;    841 T;

  Query Match             4.1%;  Score 59;  DB 11;  Length 2992;
  Best Local Similarity  64.8%;  Pred. No. 2.63e-22;
  Matches  129;  Conservative    0;  Mismatches  70;  Indels   0;  Gaps   0;

Db    561 ctgcttcgtctgccaggacaaatcatcagggtaccactatggggtcagcgcctgtgaggg 620
          ||||   ||||||   ||  |  |||| || || |||||||| ||| |  | ||||| ||
Qy    285 CTGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGG 344

Db    621 atgtaagggctttttccgcagaagtattcagaagaatatgatttacacttgtcacccaga 680
          |||||||| ||||||    ||||| |||||     |||    ||| | |||||   |
Qy    345 ATGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTAC 404

Db    681 taagaactgtgttattaataaagtcaccaggaatcgatgccaatactgtcgactccagaa 740
           ||  | |||   ||  |||||  |    | ||  | |||||   ||| ||||| | |||
Qy    405 AAATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAA 464

Db    741 gtgctttgaagtgggaatg 759
          ||| |  ||||||||||||
Qy    465 GTGTTACGAAGTGGGAATG 483


Search completed: Thu Dec 18 18:00:31 1997
Job time : 165 secs.
```

```
*******************************************************************

                 MPSRCH (TM)

*******************************************************************

          Release 2.1D John F. Collins, Biocomputing Research Unit
          Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
                  Distribution rights by IntelliGenetics, Inc.

MPsrch_nn   n.a. - n.a. database search, using Smith-Waterman algorithm

Run on:          Thu Dec 18 13:46:20 1997   MasPar time 141.99 Seconds
                                            977.491 Million cell updates/sec
Tabular output not generated.

Title:           >US-08-826-361A-20
Description:     (1-1257) from US08826361A.seq
Perfect Score:   1257
N.A. Sequence:      1 ATGAATTACAGCATTCCCAG......GGCATGCGAGGTCTGCCTGA 1257
    Comp              TACTTAATGTCGTAAGGGTC......CCGTACGCTCCAGACGGACT

Scoring table:   TABLE default
                 Gap 6

Nmatch   STD :   Dbase 0;  Query 0

Searched:        142080 seqs, 52183452 bases x 2

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database:        n-geneseq28
                   1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
                   8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
                   14:part14 15:part15 16:part16 17:part17 18:part18
                   19:part19 20:part20 21:part21 22:part22 23:part23
                   24:part24 25:part25 26:part26 27:part27 28:part28
                   29:part29

Statistics:      Mean 9.225;  Variance 5.124;  scale 1.801

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES
                  %
Result          Query
   No.  Score   Match Length DB  ID               Description            Pred. No.
----------------------------------------------------------------------------------
     1   1247   99.2   1460 29  T62843           Human oestrogen recep  0.00e+00
     2    861   68.5   2058 29  T62842           Rat oestrogen recepto  0.00e+00
     3    843   67.1   1458 29  T62944           Mouse oestrogen recep  0.00e+00
     4    135   10.7   2992  2  N70990           cDNA encoding human o  2.05e-76
     5    135   10.7   4963 17  T06973           PhCMV*-1 promoter and  2.05e-76
     6    135   10.7   4963 13  Q76270           PhCMV*-1-controlled p  2.05e-76
     7    127   10.1    204 14  Q85976           Human estrogen recept  1.43e-70
     8     91    7.2   1047  2  Q10572           Human Natriuretic Pep  1.14e-44
c    9     71    5.6   3460 24  T42193           Fragment of pig oestr  1.27e-30
    10     68    5.4   2153  2  N80922           Sequence encoding hum  1.21e-28
c   11     66    5.3   1047  2  Q10572           Human Natriuretic Pep  2.80e-27
    12     63    5.0   6244 13  Q76269           PhCMV*-1-controlled p  3.03e-25
    13     63    5.0   6244 17  T06972           PhCMV*-1 promoter and  3.03e-25
    14     59    4.7   2992 13  Q79934           Human liver hap cDNA.  1.47e-22
    15     59    4.7   2992 11  Q65572           Human liver hap cDNA.  1.47e-22
    16     59    4.7   2992 20  T06491           Human hepatoma retino  1.47e-22
    17     58    4.6   1912  2  Q10388           Encodes skin specific  5.85e-22
    18     58    4.6   2521  3  Q20542           Mouse retinoic acid r  5.85e-22
    19     58    4.6   2749  3  Q20541           Mouse retinoic acid r  5.85e-22
    20     58    4.6   5043  2  N80916           Sequence encoding the  5.85e-22
    21     55    4.5   1911 15  T15232           Murine retinoic acid   1.45e-20
    22     55    4.5   6176 27  T43226           Plasmid pGRO4... for   1.45e-20
    23     55    4.4   2099  1  N90093           Human Hap (hepatoma)   6.69e-20
    24     53    4.2   1231  5  [illegible]      Sequence [illegible]   1.39e-18
    25     53    4.2   1795  2  Q12005           Human TR2-9 DNA bindi  1.39e-18
    26     53    4.2   2029  2  Q12004           Human TR2-5 androgen   1.39e-18
    27     53    4.2   2221  2  Q12006           Human TR2-11 DNA bind  1.39e-18
    28     53    4.2   2458  2  Q12003           Human TR2-7 DNA bindi  1.39e-18
    29     51    4.1   1576  2  T10226           Clone pGEM-hRARgamma   2.80e-17
    30     51    4.1   1423  2  [illegible]      Encodes human Retinoi  2.80e-17
    31     51    4.1   1903 18  T05196           Human retinoic acid r  2.80e-17
    32     51    4.1   3569  1  N91772           Human androgen recept  2.80e-17
    33     51    4.1   3569 29  T63407           Androgen receptor cDN  2.80e-17
    34     51    4.1   3715  2  Q12001           Full-length human and  2.80e-17
    35     50    4.0   1351 13  Q79935           Human liver RAR-beta   1.25e-16
    36     50    4.0   1143  1  [illegible]      Sequence encoding hum  [illegible]
c   37     49    3.9    204  1  N81164           Base substituted E.co  5.55e-16
    38     47    3.7    387 19  T09963           Modified human glucoc  1.07e-14
    39     47    3.7    387 24  T45454           Modified human glucoc  1.07e-14
    40     47    3.7   3217  1  N91578           Rat androgen receptor  1.07e-14
    41     45    3.6   1622  1  [illegible]      TrpE/androgen recepto  2.01e-13
    42     45    3.6   3217  2  Q12002           Full-length rat andro  2.01e-13
    43     45    3.6   3427 10  [illegible]      cDNA sequence encodin  2.01e-13
    44     45    3.6   4180  1  N91773           Rat androgen receptor  2.01e-13
    45     44    3.5   4400 18  T16151           Apoptopic cerebral ne  8.60e-13

                              ALIGNMENTS

RESULT     1
ID   T62843 standard; cDNA; 1460 BP.
AC   T62843;
DT   08-JUN-1997  (first entry)
DE   Human oestrogen receptor beta (ER-beta) cDNA.
KW   Orphan receptor; oestrogen receptor beta; ER-beta;
KW   nuclear receptor; prostate cancer; benign prostatic hyperplasia;
KW   ovary cancer; cardiovascular disease; osteoporosis;
KW   environment; pollutant; ss.
OS   Homo sapiens.
FH   Key              Location/Qualifiers
FT   CDS              3..1460
FT   /*tag=  a
PN   WO9709348-A2.
PD   13-MAR-1997.
PF   09-SEP-1996; E03933.
PR   08-SEP-1995; GB-018272.
PR   15-MAR-1996; GB-005550.
PR   11-APR-1996; GB-007532.
PR   08-MAY-1996; GB-009576.
PA   (KARO-) KARO BIO AB.
PI   Enmark E,  Gustafsson J,  Kuiper GG;
DR   WPI; 97-192842/17.
DR   P-PSDB; W14724.
PT   New isolated oestrogen receptor beta - used to develop drugs for
PT   treating e.g. cancers, CNS diseases, osteoporosis or cardiovascular
PT   disease
PS   Claim 6; Fig 13B; 45pp; English.
CC   A cDNA clone (T62843) codes for a novel human oestrogen receptor-
CC   related nuclear receptor (W14724), designated ER-beta.  It was
CC   isolated from a human testis cDNA library using rat ER-beta cDNA
CC   (see also T62842) as probe.  The human ER-beta gene was mapped
CC   to chromosome 14 region 14q22-23.  Rat, human and mouse ER-beta
CC   (W14723-25) can be used to isolate molecules for use in the
CC   treatment of cardiovascular diseases, central nervous system
CC   diseases, osteoporosis, prostate or ovarian cancer or benign
CC   prostatic hyperplasia and to test environmental chemicals for
CC   oestrogenic activity.
SQ   Sequence   1460 BP;    352 A;    379 C;    411 G;    318 T;
```

```
  Query Match            99.2%;  Score 1247;  DB 29;  Length 1460;
  Best Local Similarity  100.0%;  Pred. No. 0.00e+00;
  Matches 1247;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Db      27 atgaattacagcattcccagcaatgtcactaacttggaaggtgggcctggtcggcagacc 86
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1 ATGAATTACAGCATTCCCAGCAATGTCACTAACTTGGAAGGTGGGCCTGGTCGGCAGACC 60

Db      87 acaagcccaaatgtgttgtggccaacacctgggcacctttctcctttagtggtccatcgc 146
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      61 ACAAGCCCAAATGTGTTGTGGCCAACACCTGGGCACCTTTCTCCTTTAGTGGTCCATCGC 120

Db     147 cagttatcacatctgtatgcggaacctcaaaagagtccctggtgtgaagcaagatcgcta 206
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     121 CAGTTATCACATCTGTATGCGGAACCTCAAAAGAGTCCCTGGTGTGAAGCAAGATCGCTA 180

Db     207 gaacacaccttacctgtaaacagagacactgaaaaggaaggttagtgggaaccgttgc 266
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     181 GAACACACCTTACCTGTAAACAGAGACACTGAAAAGGAAGGTTAGTGGGAACCGTTGC 240

Db     267 gccagccctgttactggtccaggttcaaagagggatgctcacttctgcgctgtctgcagc 326
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     241 GCCAGCCCTGTTACTGGTCCAGGTTCAAAGAGGGATGCTCACTTCTGCGCTGTCTGCAGC 300

Db     327 gattacgcatcgggatatcactatggagtctggtcgtgtgaaggatgtaaggcctttttt 386
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     301 GATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATGTAAGGCCTTTTTT 360

Db     387 aaaagaagcattcaaggacataatgattatatttgtccagctacaaatcagtgtacaatc 446
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     361 AAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCAGTGTACAATC 420

Db     447 gataaaaaccggcgcaagagctgccaggcctgccgacttcggaagtgttacgaagtggga 506
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     421 GATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTACGAAGTGGGA 480

Db     507 atggtgaagtgtggctcccggagagagagatgtgggtaccgccttgtgcggagacagaga 566
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     481 ATGGTGAAGTGTGGCTCCCGGAGAGAGAGATGTGGGTACCGCCTTGTGCGGAGACAGAGA 540

Db     567 aatgccgacgagcagctgcactgtgccggcaaggccaagagaagtggcggccacgcgccc 626
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     541 AATGCCGACGAGCAGCTGCACTGTGCCGGCAAGGCCAAGAGAAGTGGCGGCCACGCGCCC 600

Db     627 cgagtgcgggagctgctgctggacgccctgagccccgagcagctagtgctcaccctcctg 686
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     601 CGAGTGCGGGAGCTGCTGCTGGACGCCCTGAGCCCCGAGCAGCTAGTGCTCACCCTCCTG 660

Db     687 gaggctgagccgccccatgtgctgatcagccgccccagtgcgcccttcaccgaggcctcc 746
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     661 GAGGCTGAGCCGCCCCATGTGCTGATCAGCCGCCCCAGTGCGCCCTTCACCGAGGCCTCC 720

Db     747 atgatgatgtccctgaccaagttggccgacaaggagttggtacacatgatcagctgggcc 806
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     721 ATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTACACATGATCAGCTGGGCC 780

Db     807 aagaagattcccggctttgtggagctcagcctgttcgaccaagtgcggctcttggagagc 866
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     781 AAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAAGTGCGGCTCTTGGAGAGC 840

Db     867 tgttggatggaggtgttaatgatggggctgatgtggcgctcaattgaccaccccggcaag 926
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     841 TGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCAATTGACCACCCCGGCAAG 900

Db     927 ctcatctttgctccagatcttgttctggacagggatgaggggaaatgcgtagaaggaatt 986
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     901 CTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGGAAATGCGTAGAAGGAATT 960

Db     987 ctggaaatctttgacatgctcctggcaactacttcaaggtttcgagagttaaaactccaa 1046
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     961 CTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTTCGAGAGTTAAAACTCCAA 1020

Db    1047 cacaaagaatatctctgtgtcaaggccatgatcctgctcaattccagtatgtaccctctg 1106
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    1021 CACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAATTCCAGTATGTACCCTCTG 1080

Db    1107 gtcacagcgacccaggatgctgacagcagccggaagctggctcacttgctgaacgccgtg 1166
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    1081 GTCACAGCGACCCAGGATGCTGACAGCAGCCGGAAGCTGGCTCACTTGCTGAACGCCGTG 1140

Db    1167 accgatgctttggtttgggtgattgccaagagcggcatctcctcccagcagcaatccatg 1226
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    1141 ACCGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATCTCCTCCCAGCAGCAATCCATG 1200

Db    1227 cgcctggctaacctcctcatgctcctgtcccacgtcaggcatgcgag 1273
           |||||||||||||||||||||||||||||||||||||||||||||||
Qy    1201 CGCCTGGCTAACCTCCTCATGCTCCTGTCCCACGTCAGGCATGCGAG 1247


RESULT    2
ID   T62842 standard; cDNA; 2568 BP.
AC   T62842;
DT   08-JUN-1997  (first entry)
DE   Rat oestrogen receptor beta (ER-beta) cDNA.
KW   Orphan receptor; oestrogen receptor beta; ER-beta;
KW   nuclear receptor; prostate cancer; benign prostatic hyperplasia;
KW   ovary cancer, cardiovascular disease, osteoporosis;
KW   environment, pollutant, ss.
OS   Rattus sp.
FH   Key               Location/Qualifiers
FT   CDS               424..1881
FT   /*tag=  a
FT   /note= "alternative start codon at 448..450"
PN   WO9709348-A2.
PD   13-MAR-1997.
PF   09-SEP-1996; E03933.
PR   08-SEP-1995; GB-018272.
PR   15-MAR-1996; GB-005550.
PR   11-APR-1996; GB-007532.
PR   08-MAY-1996; GB-009576.
PA   (KARO-) KARO BIO AB.
PI   Enmark E,  Gustafsson J,  Kuiper GG;
DR   WPI; 97-192842/17.
DR   P-PSDB, W14723.
PT   New isolated oestrogen receptor beta - used to develop prods. for
PT   treating e.g. cancers, CNS diseases, osteoporosis or cardiovascular
PT   disease
PS   Claim 6; Fig 1; 45pp; English.
CC   A cDNA clone (T62842) codes for a novel rat oestrogen receptor-
CC   related nuclear receptor (W13723), designated ER-beta.  A set of
CC   degenerate primers based on conserved sequences of the DNA-
CC   binding domain and ligand binding domain of nuclear receptors were
CC   used to amplify rat prostate cDNA.  The PCR prod. was used to
CC   screen a rat prostate cDNA library to obtain the ER-beta cDNA.
CC   This was then used to obtain human ER-beta cDNA (see also T62843)
CC   from a human testis cDNA library.  Rat, human and mouse ER-beta
CC   (W14723-25) can be used to isolate molecules for use in the
CC   treatment of cardiovascular diseases, central nervous system
CC   diseases, osteoporosis, prostate or ovarian cancer or benign
CC   prostatic hyperplasia and to test environmental chemicals for
CC   oestrogenic activity.
SQ   Sequence   2568 BP;    634 A;    654 C;    663 G;    617 T;

  Query Match            68.5%;  Score 861;  DB 29;  Length 2568;
  Best Local Similarity  84.7%;  Pred. No. 0.00e+00;
  Matches 1051;  Conservative    0;  Mismatches 190;  Indels    0;  Gaps    0;

Db     448 atgaactacagtgttcccggcagcaccagtaacctggacggtggcctgtccactgagc 507
           ||||| |||||  ||||| |||     || |||| |||| ||||||||||   || | || |
Qy       1 ATGAATTACAGCATTCCCAGCAATGTCACTAACTTGGAAGGTGGGCCTGGTCGGCAGACC 60

Db     508 acaagcccaaatgtgctatggccaacttctgggcacctatctcctttagcgacccattcc 567
           ||||||||||||||| | ||||||||   |||||||||| | |||||||| |    || |   |
Qy      61 ACAAGCCCAAATGTGTTGTGGCCAACACCTGGGCACCTTTCTCCTTTAGTGGTCCATCGC 120
```

```
Db     568 caatcatcgctcctctatgcagaacctcaaaagagtccttggtgtgaagcaagatcacta 627
           || |  |||  |    || |||||  |||||||||||||||| ||||||||||||||| |||
Qy     121 CAGTTATCACATCTGTATGCGGAACCTCAAAAGAGTCCCTGGTGTGAAGCAAGATCGCTA 180

Db     628 gagcacaccttacctgtaaacagagagacactgaagaggaagcttagtgggagcagttgt 687
           || |||||||||||||| ||||||||||||||||:||| |||||| |||||||||| | ||||
Qy     181 GAACACACCTTACCTGTAAACAGAGAGACACTGAAAAGGAAGGTTAGTGGGAACCGTTGC 240

Db     688 gccagccctgttactagtccaaacgcaaagagggatgctcacttctgccccgtctgcagc 747
           |||||||||||||||| |||||     |||||||||||||||||||||||| | ||||||||
Qy     241 GCCAGCCCTGTTACTGGTCCAGGTTCAAAGAGGGATGCTCACTTCTGCCCTGTCTGCAGC 300

Db     748 gattatgcatctgggtatcattacggcgtttggtcatgtgaaggatgtaaggcctttttt 807
           ||||| |||||  |  |||| || ||  || |||| ||||||||||||||||||||||||
Qy     301 GATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATGTAAGGCCTTTTTT 360

Db     808 aaaagaagcattcaaggacataatgattatatctgtccagccacgaatcagtgtacgata 867
           |||||||||||||||||||||||||||||||| ||||||| || |||||||||||| ||
Qy     361 AAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCAGTGTACAATC 420

Db     868 gacaagaacaggcgtaaaagttgccaggcctgccgactttgcaagtgttatgaagtagga 927
           || || |||||||| || |||||||||||||||||||||||| |||||||| ||||| |||
Qy     421 GATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTACGAAGTGGGA 480

Db     928 ctggtcaagtgtggatccaggagagaacggtgtgggtaccgtatagtgcggaggcagaga 987
           :|||| ||||||||| ||| |||||||    | |||||||:||||   | |||||||| ||||||
Qy     481 ATGGTGAAGTGTGGCTCCCGGAGAGAGAGATGTGGGTACCGCCTTGTGCGGAGACAGAGA 540

Db     988 agttctagcgagcaggtacactgcctgagcaaagccaagagaaacggtgggcatgcaccc 1047
           ||| |   ||||||| | |||||      |||| ||||||||||  || || || || |||
Qy     541 AGTCCGACGAGCAGCTGCATTGTCGCGGCAAGCCAAGAGAATCGGCGGCAGCAGCGCCC 600

Db    1048 ccggtgaaggagctactgctggacgccttgagtccagagcagttggtgctcaccctcctg 1107
           || |||  ||||'| ||||||  | || |||| || ||||| || ||||||||||||||||
Qy     601 CGAGTGCGGGAGCTGCTGCTGGACGCCCTGAGCCCCGAGCAGCTAGTGCTCACCCTCCTG 660

Db    1108 gaagctgaaccacccaatgtgctggtgagccgtcccagcatgcccttcaccgaggcctcc 1167
           || ||||| || ||| |||||||| | ||||| |||||    ||||||||||||||||||||
Qy     661 GAGGCTGAGCCGCCCCATGTGCTGATCAGCCGCCCCAGTGCGCCCTTCACCGAGGCCTCC 720

Db    1168 atgatgatgtccctcactaagctggcggacaaggaactggtgcacatgattggctgggcc 1227
           |||||||||||||| || ||| |||| ||||||||  |||| ||||||||  ||||||||
Qy     721 ATGATGATGTCCCTGACCAAGTTGGCCGACAAGGAGTTGGTACACATGATCAGCTGGGCC 780

Db    1228 aagaaaatccctggctttgtggagctcagcctgttggaccaagtccggctcttagaaagc 1287
           ||||| || || ||||||||||||||||||||||| |||||||| |||||||| || |||
Qy     781 AAGAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAAGTGCGGCTCTTGGAGAGC 840

Db    1288 tgctggatggagatgctaatgatgggactgatgtggcgctcaattgaccacccgggcaag 1347
           || |||||||||||| ||||| |||| ||||||||||||||| || ||||||||||||||
Qy     841 TGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCAATTGACCACCCCGGCAAG 900

Db    1348 ctcattttcgctcccgacctcgttctggacagggatgagggaaagtgcgtagaggggatt 1407
           ||||| || |||| || || |||||||||||||||||||||||||| |||||||||| |||
Qy     901 CTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGGAAATGCGTAGAAGGAATT 960

Db    1408 ctggaaatctttgacatgctcctggcgagtacgtcaaggttccgtgagttaaaactccag 1467
           |||||||||||||||||||||||||| || || |||||||| || |||||||||||||
Qy     961 CTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTTCGAGAGTTAAAACTCCAA 1020

Db    1468 cacaaggaatatctctgtgtcaaggccatgatcctgctcaattccagtatgtaccctttg 1527
           ||||| || ||||||||||| |||||||||||||| ||||| |||||||||||||||  ||
Qy    1021 CACAAAGAATATCTCTGTGTCAAGGCCATGATCCTGCTCAATTCCAGTATGTACCCTCTG 1080

Db    1528 gcttctgcagcccaggaagcagaaaatagccagaagctgcacacttactgaacgcggtg 1587
           |   |  ||| |  |||||| || || ||  ||||||||||||| |||| | ||||||| |||
Qy    1081 GTCACAGCGACCCAGGATGCTGACAGCAGCCGGAAGCTGGCTCACTTGCTGAACGCCGTG 1140

Db    1588 acagatgccctggtctgggtgattgcgaagagtggtatctcctcccagcagcagtcagtc 1647
           || |||||   |||| ||||||||||| ||||| || ||||||||||||||||| ||   |
Qy    1141 ACCGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATCTCCTCCCAGCAGCAATCCATG 1200

Db    1648 cgactggccaacctcctaatgcttctttctcacgtcaggca 1688
           || ||||| ||||||||:||||||||| || ||||||||'|.
Qy    1201 CGGCTGGCTAACCTCCTGATGCTCCTGTCCCACGTCAGGCA 1241

RESULT    3
ID   T62844 standard; cDNA; 1458 BP.
AC   T62844;
DT   08-JUN-1997  (first entry)
DE   Mouse oestrogen receptor beta (ER-beta) cDNA.
KW   Orphan receptor; oestrogen receptor beta; ER-beta;
KW   nuclear receptor, prostate cancer, benign prostatic hyperplasia;
KW   ovary cancer, cardiovascular disease, osteoporosis;
KW   environment; pollutant; ss.
OS   Mus sp.
PN   WO9709348-A2.
PD   13-MAR-1997.
PF   09-SEP-1996; E03933
PR   08-SEP-1995; GB-018272.
PR   15-MAR-1996; GB-005550.
PR   11-APR-1996; GB-007532.
PR   08-MAY-1996; GB-009576.
PA   (KARO-) KARO BIO AB.
PI   Enmark E, Gustafsson J, Kuiper GG;
DR   WPI; 97-192842/17.
DR   P-PSDB; W14725.
PT   New isolated oestrogen receptor beta - used to develop prods. for
PT   treating e.g. cancers, CNS diseases, osteoporosis or cardiovascular
PT   disease
PS   Claim 6; Fig 14B; 45pp; English.
CC   A cDNA clone (T62844) codes for a novel mouse oestrogen receptor-
CC   related nuclear receptor (W13725), designated ER-beta.  Rat, human
CC   and mouse ER-beta (W14723-25) can be used to isolate molecules for
CC   use in the treatment of cardiovascular diseases, central nervous
CC   system diseases, osteoporosis, prostate or ovarian cancer or benign
CC   prostatic hyperplasia and to test environmental chemicals for
CC   oestrogenic activity.
SQ   Sequence   1458 BP;    365 A;    373 C;    403 G;    317 T;

  Query Match             67.1%,  Score 842,  DB 23,  Length 1458;
  Best Local Similarity   84.0%;  Pred. No. 0.00e+00;
  Matches  1042,  Conservative    0,  Mismatches 199,  Indels    0,  Gaps    0;

Db      25 atgaactacagtgttcccagcagcaccggtaacctggaaggtgggcctgtccgcagact 84
           ||||| |||||  |||||||||      |   |||| ||||||||||||||| ||| |||||
Qy       1 ATGAATTACAGCATTCCTAGCAATGTCACTAACTTGGAAGGTGGGCCTGGTCGGCAGACC 60

Db      85 gcaagcccaaatgtgctatggccaattctggacacctctctcctttagcccacccactgc 144
            |||||||||||||| | |'||||||   |||| |||||  |'||||'||||    |||  ||
Qy      61 ACAAGTCCAAATGTGTTGTTGGCAATACTTGTGCACCTTTTTCCTTTATTGGTTCATTCC 120

Db     145 caatcatcgcttctctatgcagaacctcaaaagagtccttggtgtgaagcaagatcacta 204
           || | ||| | ||| ||||| ||||||||||||||||| ||||||||||||,'|||| |||
Qy     121 CAGTTATCACATCTGTATGCGGAACCTCAAAAGAGTCCCTGGTGTGAAGCAAGATCGCTA 180

Db     205 gaacacaccttgcctgtaaacagagagaccctgaagaggaagcttggccggagccgttgt 264
           ||||||||||| ||||||||||||||||| ||||| |||||| || | |||| | ||||
Qy     181 GAACACACCTTACCTGTAAACAGAGAGACACTGAAAAGGAAGGTTAGTGGGAACCGTTGC 240

Db     265 gccagccctgttactagtccaagcaccaagagggatgctcacttctgtgccatctgcagt 324
           ||||||||||||||| ||||| |   | |||||||||||||:|||||||| || ||||||'
Qy     241 GCCAGCCCTGTTACTGGTCCAGGTTCAAAGAGGGATGCTCACTTCTGCCCTGTCTGCAGC 300

Db     325 gattatgcatctgggtatcattacggtgtctggtcctgtgaaggatgtaaggcctttttt 384
           ||||| ||||| || ||||| || || |||||||| ||||||||||||||||||||||||
Qy     301 GATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATGTAAGGCCTTTTTT 360

Db     385 aaaagaagcattcaaggacataatgattatatctgtccagccacgaatcagtgtacgata 444
           |||||||||||||||||||||||||| ||||| |||||||| || ||||:|||||| ||
Qy     361 AAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCAGTGTACAATC 420

Db     445 gacaagaaccggcgtaaaaactgccaggcctgccgactttgcaagtgttacgaaatagga 504
           || || |||||||| || | |||||||||||||||||||| ||||||||'|'||:| |:|
Qy     421 GATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTACGAAGTGGGA 480
```

```
Db     505 atggtcaagtgtggatccaggagaaaaggtgtgggtaccgaatagtacgaagacagaga 564
           ||||| |||||||| ||| ||||||| || |||||||||||   | || || |||||||||
Qy     481 ATGGTGAAGTGTGGCTCCCGGAGAGAGAGATGTGGGTACCGGCTTGTGCGGAGACAGAGA 540

Db     565 agtgccagcgagcaggtgcattgcctgaacaaatccaagaggaccagtggcacatacctc 624
           ||||||  ||||||| |||| ||        ||| ||||||||||    | || ||| | |||
Qy     541 AGTGCCGACGAGCAGCTGCACTGTCCGGCAAGGCCAAGAGAAGTGGCGGCCACGCGCCC 600

Db     625 caagtgaagcaactaatcctctcaacccttaagtcccgagcagctggtgctcaccctgctg 684
           || |||   |||||| |||||| || | ||||| ||||||||||| |||||||||| |||
Qy     601 CGAGTGCGGCAGCTCCTGCTGGACCCCCTGAGCCCGAGCAGCTAGTGCTCACCCTCCTG 660

Db     685 gaagctgagccacccaatgtgctagtgagtcgtcccagcatgcccttcaccgaggcctcc 744
           || |||||||| ||| |||||||    | || || |||||    |||||||||||||||||||
Qy     661 GAGGCTGAGCCGCCCCATGTGCTGATCAGCCGCCCCAGTGCGCCCTTCACCGAGGCCTCC 720

Db     745 atgatgatgtcccttaccaagctggctgacaaggaactggtgcacatgattggctgggcc 804
           |||||||||||||| || ||| |||| ||||||||   |||| ||||||||   ||||||||
Qy     721 ATGATGATGTCCCTGACCAACTTGCCCACAAGGCACTTCGTACACATGATCAGCTGGGCC 780

Db     805 aagaaaatccctggctttgtggagctcagcctgttggaccaagtccggctcttggaaagc 864
           ||||| || || |||||||||||||||||||||| ||||||||| || ||||||||| |||
Qy     781 AACAAGATTCCCGGCTTTGTGGAGCTCAGCCTGTTCGACCAAGTCCGCCTCTTGGAGAGC 840

Db     865 tgctggatggaggtgttaatgatggggctgatgtggcgctccatcgaccaccccggcaag 924
           || ||||||||||||| | ||| |||||||||||||||||||| || |||||||||||||||
Qy     841 TGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGGCGCTCAATTGACCACCCCGGCAAG 900

Db     925 ctcatctttgctccagacctgattctagacaggatgaggggaatgcatgaaaggatt 984
           ||||||||||||||||| || |||||||||||||||||||||||| ||||| ||||| |||
Qy     901 CTCATCTTTGCTCCAGATCTTGTTCTGGACAGGGATGAGGGGAAATGCGTAGAAGGAATT 960

Db     985 ctggaaatctttgacatgctcctggcaactacttcaaggttccgtgagttaaaactgcag 1044
           ||||||||||||||||||||||||| || ||   || |||| || ||||||||||| ||
Qy     961 CTGGAAATCTTTGACATGCTCCTGGCAACTACTTCAAGGTTCGAGAGTTAAAACTCCAA 1020

Db    1045 cacaaagaatatctgtgtgtgaaggccatgattctcctcaactccagtatgtaccacttg 1104
           |||||||||||||| ||||| |||||||||||| || ||||| |||||||||||||    ||
Qy    1021 CACAAAGAATATCTGTGTGTCAAGGCCATGATCCTGCTCAATTCCAGTATGTACCCTCTG 1080

Db    1105 gctaccgcaagccaggaagcagagagtagccggaagctgacacacctattgaacgcagtg 1164
           |  || || | |||||| || || || ||||||||||||| | ||| |   ||||||| |||
Qy    1081 GTCACAGCGACCCAGGATGCTGACAGCAGCCGGAAGCTGGCTCACTTGCTGAACGCCGTG 1140

Db    1165 acagatgccctggtctgggtgatttcgaagagtagaatctcttcccagcagcagtcagtc 1224
           || |||||   |||| |||||||||| | |||||   | ||||| ||||||||||| ||  |
Qy    1141 ACCGATGCTTTGGTTTGGGTGATTGCCAAGAGCGGCATCTCCTCCCAGCAGCAATCCATG 1200

Db    1225 cgtctggccaacctcctgatgcttctttctcatgtcaggca 1265
           || ||||| ||||||||||||||||| || || || ||||||||
Qy    1201 CGCCTGGCTAACCTCCTGATGCTCCTGTCCCACGTCAGGCA 1241


RESULT    4
ID   N70880 standard; cDNA; 2092 BP.
AC   N70880;
DT   24-FEB-1991  (first entry)
DE   cDNA encoding human oestrogen receptor protein
KW   Human oestrogen receptor protein; ss cDNA; expression system; assay;
KW   steroid receptor protein;
OS   Homo sapiens.
FH   Key                 Location/Qualifiers
FT   CDS                 293..1290
FT   /*tag= a
PN   WO8705049-A.
PD   27-AUG-1987.
PF   18-FEB-1987; U00341.
PR   20-FEB-1986; US-833829.
PA   (CALB-) California Biotechnology Inc
PI   Shine J;
DR   WPI; 87-250213/35.
DR   P-PSDB; P70542.
PT   Expression system for vertebrate steroid receptor protein - comprising
PT   DNA sequence encoding the protein linked to control sequences in
PT   eucaryotic hosts.
PS   Disclosure; Fig. 1-1 - 1-2, 26pp; English.
CC   The cDNA can be used in an expression system to express human
CC   oestrogen receptor protein.  The coding sequence is operably linked to
CC   control sequences compatible with eukaryotic host cells.  This method
CC   allows expression under conditions which favour appropriate post-
CC   translational processing.  It produces large amts. of purified protein
CC   useful in the design of agonist and antagonist cpds., for the study of
CC   the mechanism of action of the steroid binding proteins in general,
CC   and for use in diagnostic assays for the proteins or antibodies to
CC   them.  These assays are important in, eg the diagnosis of tumour
CC   sensitivities to steroid metabolism.  Suitable host cells are VERO,
CC   HeLa and CHO cells.
SQ   Sequence   2092 BP;    473 A;    607 C;    591 G;    421 T;

  Query Match              10.7%;  Score 135;  DB 2;  Length 2092;
  Best Local Similarity  77.8%;  Pred. No. 2.05e-76;
  Matches  189;  Conservative    0;  Mismatches  54;  Indels    0;  Gaps    0;

Db     832 ggagactcgctactgtgcagtgtgcaatgactatgcttcaggctaccattatggagtctg 891
            |||   ||| || ||| || || ||||   || || || || || || || ||||||||||
Qy     273 GGATGCTCACTTCTGGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTG 332

Db     892 gtcctgtgagggctgcaaggcctttttcaagagaagtattcaaggacataaccgactatat 951
           ||| ||||| || || |||||||| || || ||||| ||||||||||||| || |||||
Qy     333 GTCGTGTGAAGGATGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATAT 392

Db     952 gtgtccagccaccaaccagtgcaccattgataaaaacaggaggaagagctgccaggcctg 1011
            |||||||| || || ||||| || || ||||||||| || | |||||||||||||||||
Qy     393 TTGTCCAGCTACAAATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTG 452

Db    1012 ccggctccgcaagtgctacgaagtgggaatgatgaaaggtgggatacggaaagaccgaag 1071
           ||| || || || || ||||||||||||||| ||||  ||||      || | ||| || |
Qy     453 CCGACTTCGGAAGTGTTACGAAGTGGGAATGGTGAAGTGTGGCTCCCGGAGAGAGAGATG 512

Db    1072 agg 1074
            ||
Qy     513 TGG 515


RESULT    5
ID   T06873 standard; DNA; 4963 BP.
AC   T06873;
DT   13-MAR-1996  (first entry)
DE   PhCMV*-1 promoter and human oestrogen receptor.
KW   Transactivator; tTA; tet operator; tetO; gene expression;
KW   tetracycline-responsive promoter; PhCMV*-1; human cytomegalovirus;
KW   transcription activation; eukaryotic cell; oestrogen receptor;
KW   ds; ss.
OS   Chimeric synthetic;
OS   Chimeric Homo sapiens.
PN   US5464758-A.
PD   07-NOV-1995.
PF   14-JUN-1993; 076726.
PR   14-JUN-1993; US-076726.
PA   (BUJA/) BUJARD H.
PA   (GOSS/) GOSSEN M
PI   Bujard H, Gossen M;
DR   WPI; 95-392612/50.
PT   Polynucleotide encoding trans:activator fusion protein contg. tet
PT   repressor  - used to control expression of gene regulated by minimal
PT   promoter linked to tet operon, and vectors and cells where gene
PT   expression is regulated by tetracycline
PS   Disclosure; Fig 10A-B; 37pp; English.
CC   cDNA coding for the human oestrogen receptor was placed under
CC   control of the PhCMV*-1 promoter (see T06869), resulting in
CC   tetracycline-regulated expression vector pUHDpgr-3 contg. the
CC   sequence given in T06873.  The construct was transfected into
CC   eukaryotic cells expressing the tTA transactivator (T06867) to
CC   allow tetracycline-regulated expression of the receptor.
SQ   Sequence   4963 BP;    1259 A;    1300 C;    1278 G;    1126 T;
```

```
  Query Match             10.7%;  Score 135;  DB 17;  Length 4963;
  Best Local Similarity   77.8%;  Pred. No. 2.05e-76;
  Matches  189;  Conservative    0;  Mismatches   54;  Indels    0;  Gaps    0;

Db   1005 ggagactcgctactgtgcagtgtgcaatgactatgcttcaggctaccattatggagtctg 1064
          |||  ||| || ||| || || ||||  || || || || || || || |||||||||||
Qy    273 GGATGCTCACTTCTGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTG 332

Db   1065 gtcctgtgagggctgcaaggccttcttcaagagaagtattcaaggacataacgactatat 1124
          ||| ||||| || || |||||||| || || ||||| |||||||||||||| || |||||
Qy    333 GTCGTGTGAAGGATGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATAT 392

Db   1125 gtatccagccaccaaccagtgcaccattgataaaaacaggaggaagagctgccaggcctg 1184
           | |||||| || || ||||| || || ||||||||| || | |||||||||||||||||
Qy    393 TTGTCCAGCTACAAATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTG 452

Db   1185 ccggctccgcaaatgctacgaagtgggaatgatgaaaggtgggatacgaaaagaccgaag 1244
          ||| || || || || ||||||||||||||| ||||  ||||     | | | |  || |
Qy    453 CCGACTTCGGAAGTGTTACGAAGTGGGAATGGTGAAGTGTGGCTCCGGCAGACAGAGATG 512

Db   1245 agg 1247
           ||
Qy    513 TGG 515


RESULT    6
ID   Q76270 standard; DNA; 4963 BP.
AC   Q76270;
DT   17-JUL-1995  (first entry)
DE   PhCMV*-1-controlled progesterone receptor
KW   tTA; transactivator; tetracycline-controllable transactivator;
KW   conditional inactivation; homologous recombination; gene expression;
KW   gene regulation; gene therapy; tetracycline-resistance; tetR;
KW   transgenic animal; PhCMV*-1; promoter; tetO; CMV;
KW   progesterone receptor; ds.
OS   Human cytomegalovirus Towne.
PN   WO9429442-A.
PD   22-DEC-1994.
PF   14-JUN-1994; U06734
PR   14-JUN-1993; US-076327.
PA   (BADI ) BASF AG.
PI   Bujard H,  Gossen M,  Salfeld JG,  Voss JW;
DR   WPI; 95-036472/05
PT   Regulatory systems using tetracycline-controllable transactivator
PT   (tTA)  - useful for conditional inactivation or modulation of
PT   gene expression in a host cell or animal
PS   Disclosure; Page 54-58; 103pp; English.
CC   The minimal promoter PhCMV*-1 is based on the human CMV
CC   intermediate-early promoter (with enhancer region removed), fused
CC   to a tetracycline operator, tetO.  The promoter is used to
CC   express tTA transactivator in host cells.  In clone pUHD BGF4
CC   (given in Q76269), cDNA for rabbit progesterone receptor is under
CC   the control of PhCMV*-1.
SQ   Sequence   4963 BP;    1259 A;    1298 C;    1280 G;    1126 T;

  Query Match             10.7%;  Score 135;  DB 13;  Length 4963;
  Best Local Similarity   77.8%;  Pred. No. 2.05e-76;
  Matches  189;  Conservative    0;  Mismatches   54;  Indels    0;  Gaps    0;

Db   1005 ggagactcgctactgtgcagtgtgcaatgactatgcttcaggctaccattatggagtctg 1064
          |||  ||| || ||| || || ||||  || || || || || || || |||||||||||
Qy    273 GGATGCTCACTTCTGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTG 332

Db   1065 gtcctgtgagggctgcaaggccttcttcaagagaagtattcaaggacataacgactatat 1124
          ||| ||||| || || |||||||| || || ||||| |||||||||||||| || |||||
Qy    333 GTCGTGTGAAGGATGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATAT 392

Db   1125 gtatccagccaccaaccagtgcaccattgataaaaacaggaggaagagctgccaggcctg 1184
           | |||||| || || ||||| || || ||||||||| || | |||||||||||||||||
Qy    393 TTGTCCAGCTACAAATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTG 452

Db   1185 ccggctccgcaaatgctacgaagtgggaatgatgaaaggtgggatacgaaaagaccgaag 1244
```

```
          ||| || || || || ||||||||||||||| ||||  ||||     | | | |  || |
Qy    453 CCGACTTCGGAAGTGTTACGAAGTGGGAATGGTGAAGTGTGGCTCCGGCAGACAGAGATG 512

Db   1245 agg 1247
           ||
Qy    513 TGG 515


RESULT    7
ID   Q85976 standard; cDNA; 204 BP.
AC   Q85976;
DT   31-OCT-1995  (first entry)
DE   Human estrogen receptor cDNA.
KW   nuclear receptor protein; osteosarcoma; EAR-1r; generation; growth;
KW   differentiation; human; erb A related gene; estrogen receptor; ds.
OS   Homo sapiens.
PN   J07033798-A.
PD   03-FEB-1995.
PF   20-JUL-1993; 200333.
PR   20-JUL-1993; JP-200333.
PA   (MITU ) MITSUBISHI KASEI CORP.
DR   WPI; 95-110650/15.
DR   P-PSDB; R71566.
PT   Nuclear receptor protein derived from osteoblastic osteosarcoma
PT   - encoded by the EAR-1r gene, for investigation of generation of
PT   fat soluble active substances
PS   Disclosure; Page 6; 8pp; Japanese.
CC   This cDNA encodes an estrogen receptor (a member of a receptor
CC   superfamily). It is disclosed in the patent specification. The
CC   invention concerns an EAR-1r gene (human erb-A related gene)
CC   encoding a nuclear receptor protein (R71565) derived from human
CC   osteosarcoma (HOS) cells. The expression of EAR-1r is most abundant
CC   in the brain and was detected slightly in heart and placenta. The
CC   receptor protein may be used for the investigation of generation,
CC   differentiation and growth of fat soluble physiological substances.
SQ   Sequence   204 BP;    59 A;    46 C;    55 G;    44 T;

  Query Match             10.1%;  Score 127;  DB 14;  Length 204;
  Best Local Similarity   81.3%;  Pred. No. 1.43e-70;
  Matches  165;  Conservative    0;  Mismatches   38;  Indels    0;  Gaps    0;

Db      1 tgtgcagtgtgcaatgactatgcttcaggctaccattatggagtctggtcctgtgagggc 60
          || || || ||||  || || || || || || || |||||||||||||| ||||| ||
Qy    286 TGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGA 345

Db     61 tgcaaggccttcttcaagagaagtattcaaggacataacgactatatgtatccagccacc 120
          || |||||||| || || ||||| |||||||||||||| || ||||| | |||||| ||
Qy    346 TGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACA 405

Db    121 aaccagtgcaccattgataaaaacaggaggaagagctgccaggcctgccggctccgcaaa 180
          || ||||| || || ||||||||| || | |||||||||||||||||||| || || ||
Qy    406 AATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAG 465

Db    181 tgctacgaagtgggaatgatgaa 203
          || ||||||||||||||| ||||
Qy    466 TGTTACGAAGTGGGAATGGTGAA 488


RESULT    8
ID   Q10572 standard; DNA; 1047 BP.
AC   Q10572;
DT   09-APR-1991  (first entry)
DE   Human Natriuretic Peptide Receptor B.
KW   NPRB; ANP; BNP; CNP; kidney failure; heart failure; protein kinase;
KW   hyperaldosteronism; glaucoma; guanyl cyclase.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   Peptide         1..22
FT   /label= signal sequence
FT   Protein         12
FT   /label= mature NPBR
FT   Domain          23..455
FT   /label= extracellular domain
FT   /note= "binds natriuretic peptides A,B and C]"
FT   Domain            456..456
FT   /label= transmembrane domain
FT   Domain            479..1047
FT   /label= cytoplasmic domain
FT   /note= "GC and protein kinase activity"
FT   Modified-site  24..26
FT   /label= N-glycos_site
FT   Modified site  35..37
FT   /label= N-glycos_site
FT   Modified-site  161..163
FT   /label= N-glycos_site
FT   Modified-site  195..197
FT   /label= N-glycos_site
FT   Modified-site  244..246
FT   /label= N-glycos_site
FT   Modified-site  277..279
FT   /label= N-glycos_site
FT   Modified-site  349..351
FT   /label= N-glycos_site
FT   Modified-site  600..602
FT   /label= N-glycos_site
PN   WO9100292-A.
PD   10-JAN-1991.
PF   22-JUN 1990; U03586.
PR   23-JUN-1989; US-370673
PA   (GETH ) GENENTECH INC.
PI   Chang M, Goeddel D, Lowe D,
DR   WPI; 91-036711/05.
DR   N-PSDB; Q10324
PT   Natriuretic protein receptor B   for diagnosis and treatment of
PT   kidney failure, heart failure, hyperaldosteronism, glaucoma etc
PS   Claim 3; Fig 1; 49pp, English.
CC   The sequence was derived from the DNA encoding natriuretic peptide
CC   receptor B, NPRB, having guanyl cyclase (GC) activity and protein
CC   kinase activity.  The DNA can be inserted into expression vectors
CC   for the prodn. of the protein, opt  after being mutated to produce
CC   NPRB analogues. The protein has a mol wt  of 115 kD (calculated Mr-
CC   114,952).   The protein (or variants) can be used in treatment of
CC   natriuretic peptide disorders, and also to isolate peptides using
CC   affinity chromatography.  Antibodies with affinity for NPRB can
CC   also be prepd.
SQ   Sequence   1047 BP;    87 A;    15 C;    83 G;    51 T;

  Query Match             7.2%;  Score 91;  DB 2;  Length 1047;
  Best Local Similarity   8.9%;  Pred. No. 1.14e-44;
  Matches    83;  Conservative  263;  Mismatches 572;  Indels  10;  Gaps  10;

Db      108 varnashwrnnnntaqavasqnsakndhyrtnvrtqnsanknqnnvvtnhqhnnwtaraa 167
              :|: |:: :         ||:| |  |         |
Qy      198 AAACAGAGAGACACTGAAAAGGAAGGTTAGTGGAACCGTTGCGCCAGCCCTGTTACTGG 257

Db      168 nnyndart-ddrnhyntnnqvnnanngsnnsvnhnvyarnnqnnnathnnrangrnvyn 226
               : :  | :::: :       |.  |    .  .. :  : :  |      :  : |: ::
Qy      258 TCCAGGTTCAAAGAGGGATGCTCACTTCTGCGCTGTCTGCAGCGATTACGCATCGGGATA 317

Db      227 eqnnnmnhnnnnnanrnnntn qdyvnnyndvnqnsnraqntratqrnwndnrtrnnana 285
                 :          :    |  |: :  : :: | :       | |  : : : :|    |
Qy      318 TCACTATGGAGTCTGGTCGTGTGAAGGATGTAAGGCCTTTTTTAAAAGAAGCATTCAAGG 377

Db      286 [illegible] 345
                 |:     ::       :                  : :  :     : :   |   ::|    |
Qy      378 ACATAATGATTATATTTGTCCAGCTACAAATCAGTGTACAATCGATAAAAACCGGCGCAA 437

Db      346 nvnnntnnnqqtrndqnrnvnkmnqrryhqntqrnvmdknndrntdnvnwamqdndsqdn 405
              :        ||    :  :  :  :: |:: :| |  :::::  :: |:    :|   : :: :
Qy      438 GAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTACGAAGTCGGAATCGTCAAGTGTCCGTC 497

Db      406 nnaahysqanknnwwtqrnnnwvkqannsdnnnnandnddnsndktnnstnanvanqtqn 465
                 :   ||      : ||:      ::       ::        | : :: : ::|    :   | :|     |
Qy      498 CCGGAGAGAGAGATGTGGGTACCGGCTTGTGCGGAGACAGAGAAGTGCC-GACGAGCAGC 556

Db      466 tnnmnqvs[illegible] 525
```

```
            |  :  : :     :  :      |:: :: ::            ::      :: | : :| ::
Qy      557 TGCACTCTGCCG-CCAAGCCCAAGAGAA-GTGGCGGCCACGGCGCCCGAGTGCGGGAGCT 614

Db      526 snmtahakynnnantghnkqnvvankhvnkkrnnntrnvnnnnkhmrdvnnnhntrnnqa 585
              .   .|.       ||:      ::|  :: : :     |   :         : :::
Qy      615 GCTGCTGGACGCCCTGACCCCGGAGCACCTAGTGCTCACCCTCCTGGAGGCTGAGCCGCC 674

Db      586 endnnnnenvtnyenrgsnndnnnndsnnndwmnrysnnndnvkqmannhnsnnsshqsn 645
            | :     | :  :     :         ::      :   :::     : :: :|  :  :   :  :  :
Qy      675 CCATGTCCTGATCAGCCGCCCCAGTGCGCCCTTCACCGAGGCCTCCATGATGATGCCCCT 734

Db      646 kssnqvvdsrnvnkntdyqnasn-rstanndnnnnyakknntannnsqnnnnttamna 704
             : :  ::: :  :  :   : | |:   ::||   :    | :|:             :       :
Qy      735 GACCAAGTTGGCCGACAAGGAGTTGGTACACATGATCAGCTGGGCCAAGAAGATTCCCGG 794

Db      705 advysnqnnnnnanrsqnnynnqndnsnknnvnkvrnqnrnynrnsndrtnnnnnnvnn 764
              : :: |      | ::   :  |         :  : ::      :    : : : |          :
Qy      795 CTTTGTGGAGCTCA-GCCTGTTCGACCAAGTGCGGCTCTTGGAGAGCTGTTGGATGGAGG 853

Db      765 mnrcwandnanrndnqnnkqnnrrnnknqqtsnndnnnnrmnnyannnnknvnnrtnayn 824
                 :| :   : :     :|  :    :     |                  :  :|            | :
Qy      854 TGTTAATGATGGGGCTGATGTGGCGCTCAATTGACCACCCCGGCAAGCTCATCTTTGCTC 913

Db      825 nnkrkananny nnnhsvannnkrqntvnanandsvtnynsdnvqntans-anstnmnvv 883
               :..       :     :::        ::     : | | :::|      : :   |   : |  |     :
Qy      914 CAGATCTTGTTCTGGACAGGGATGAGGGGAAATGCGTAGAAGGAATTCTGGAAATCTTTG 973

Db      884 tnnndnytcndanndnndvykvntnqdaymvvsqnnqnnqnnhannnarmanannndavss 943
                 : :  | :     :  :::::     |: :::::       :         :         ::| |  : : :
Qy      974 ACATGCTCCTGGCAACTACTTCAAGGTTTCGAGAGTTAAAACTCCAACACAAAGAATATC 1033

Db      944 nrnrhrnhdnnrn-rnqvhtqnvcaqvvqnkmnrycnnqdtvntasrmnsnqnanknhvs 1002
                ... ..   .   . |..      .|   ::     : ::       : : |  :   :       : :::
Qy     1034 TCTGTGTCAAGGCCATGATCCTGCTCAATTCCAGTATGTACCCTCTGGTCACAGCGACCC 1093

Db     1003 -sttkdandnnqnnnnnnrqdvnmkqkqkmrtyw 1035
               :                    |       :|            :
Qy     1094 AGGATGCTGACAGCAGCCGGAAGCTGGCTCACTT 1127


RESULT    9
ID   T42193 standard; DNA; 3460 BP.
AC   T42193;
DT   20-JAN-1997  (first entry)
DE   Fragment of pig oestrogen receptor gene.
KW   Pig; oestrogen receptor; exon; polymorphism; allele; RFLP; litter size;
KW   restriction fragment length polymorphism, Chinese; American; PCR; primer;
KW   amplification; polymerase chain reaction; ds.
OS   Sus scrofa.
FH   Key              Location/Qualifiers
FT   misc_feature     303..304
FT   /*tag=  a
FT   /note= "undetermined length gap missing in DNA sequence
FT   between these nucleotides"
FT   exon             534..600
FT   /*tag=  b
FT   /note= "exon 3"
FT   misc_feature     1197..1198
FT   /*tag=  c
FT   /note= "approx. 100-200 bases missing from DNA sequence
FT   between these nucleotides"
FT   primer_bind      1898..1915
FT   /*tag=  d
FT   /note= "binds primer A10RII (T42200)"
FT   primer_bind      2011..2031
FT   /*tag=  e
FT   /note= "binds primer DV028 (T42198)"
FT   primer_bind      2146..2171
FT   /*tag=  f
FT   /note= "binds primer ESR REV (T42195)"
FT   primer_bind      2536
FT   /*tag=  g
FT   primer_bind      [illegible]
FT   /*tag=  h
FT   /note= "binds primer ESRSR (T42197)"
FT   variation       2688..2693
FT   /*tag=  i
FT   /note= "polymorphic PvuII site"
FT   primer_bind     2729..2755
FT   /*tag=  j
FT   /note= "binds primer ESRSF (T42196)"
FT   primer_bind     2922..2940
FT   /*tag=  k
FT   /note= "binds primer B1F (T42201)"
FT   primer_bind     2982..3002
FT   /*tag=  l
FT   /note= "binds primer DV123 (T42199)"
FT   primer_bind     3331..3355
FT   /*tag=  m
FT   /note= "binds primer ESR FOR (T42194)"
FT   misc_feature    3122..3129
FT   /*tag=  n
FT   /note= "constant PvuII site"
PN   US5550024-A.
PD   27-AUG-1996.
PF   19-APR-1991; 687708.
PR   19-APR-1991; US-687708.
PR   19-OCT-1992; US-961919.
PR   26-SEP-1994; US-312312.
PA   (BIOT-) BIOTECHNOLOGY RES & DEV CORP.
PA   (IOWA ) UNIV IOWA STATE RES FOUND INC.
PI   Jacobson CD, Mileham AJ, Plastow GS, Rothschild MF;
PI   Tuggle CK, Vaske DA;
DR   WPI; 96-401592/40.
PT   Screening pigs for high litter size - by identifying a polymorphism
PT   in the pig oestrogen receptor gene in a DNA sample
PS   Disclosure; Column 19-24; 26pp; English.
CC   This is a fragment of the pig oestrogen receptor (ESR) gene including
CC   exon 8. The sequence covers the region contg. a polymorphic PvuII
CC   restriction enzyme site which is used as site for determining the ESR
CC   allele carried by the animal. Restriction fragment length polymorphism
CC   (RFLP) analysis of this region revealed two alleles of 4.3 and 3.7 kb.
CC   Both alleles are involved with increased pig litter size dependent on the
CC   breed of pig involved e.g. Chinese or American pigs. However pigs with
CC   the 3.7 kb allele had between 1.78-2.4 more pigs born in each litter and
CC   1.8-2.4 more pigs born alive than pigs with the 4.3 kb allele. The allele
CC   can also be determined by PCR using the primers (T42194-T42201).
SQ   Sequence  3460 BP;   1052 A;   765 C;   634 G;   1009 T;

  Query Match            5.6%;  Score 71;  DB 24;  Length 3460;
  Best Local Similarity  83.3%;  Pred. No. 1.07e-20;
  Matches   95;  Conservative   0;  Mismatches  18;  Indels   1;  Gaps   1;

Db    490 ttcatcatgcccacttcgtagcacttgcgtagccggcaggcctgacagctcttcctcctg 548
           | || ||| |||||||||||| ||||| || || |||||||| || |||||||| | || |
Cp    488 TTCACCATTCCCACTTCGTAACACTTCCGAAGTCGGCAGGCCTGGCAGCTCTTGCGCCGG 429

Db    549 ttcttatcaattgtgcactggttggtagctggacacatgtagtcattatgtcct 602
          || ||||| |||| | ||| || |||||||||| || || |||||||||||
Cp    428 TTTTTATCGATTGTACACTGATTTGTAGCTGGACAAATATAATCATTATGTCCT 375


RESULT   10
ID   N80922 standard; cDNA; 2153 BP.
AC   N80922;
DT   29-DEC-1990  (first entry)
DE   Sequence encoding human estr gen-related receptor protein (hERR2);
KW   Hormone receptor; hormone-binding; transcription activation; ss.
OS   Homo sapiens
FH   Key               Location/Qualifiers
FT   CDS               100..1401
FT   /*tag=  a
PN   WO8803168-A.
PD   05-MAY-1988.
PF   23-OCT-1987; U02782.
PR   4-OCT-1986; US-922595
```

```
PR   20-OCT-1987; US-108471.
PA   (SALK) Salk Inst for Biol Stud.
PI   Evans RM, Weinberger CA, Hollenberg SM, Giguere V;
DR   WPI; 88-133242/19.
DR   P-PSDB; P80931.
PT   Recombinant DNA encoding hormone recpetors -
PT   comprising glucocorticoid, mineralocorticoid, thyroid hormone
PT   and novel hormone receptors
PS   Claim 16; Fig V-1(B)-1 and -2; 243pp; English.
CC   DNAs encoding hormone receptors and the hormone receptors themselves are
CC   claimed. The DNA can be used to make the hormone receptor proteins and
CC   functional modified forms in quantities not previously possible. The
CC   receptor proteins can be used to screen cpds. for receptor-agonist or
CC   receptor-antagonist activity. They can also be used in diagnostic assays.
CC   Also claimed is pure DNA encoding oestrogen-related receptors hERR1 and
CC   hERR2 and hERR1 and hERR2. The new hERR1 and hERR2 receptors will provide
CC   the basis for development of an assay system that will lead to the
CC   identification of novel hormones.
SQ   Sequence  2153 BP;   445 A;   627 C;   624 G;   457 T;

  Query Match            5.4%;  Score 68;  DB 2;  Length 2153;
  Best Local Similarity  67.2%;  Pred. No. 1.21e-28;
  Matches  133;  Conservative   0;  Mismatches  65;  Indels   0;  Gaps   0;

Db    412 gtgtcgggacattgcttctggctaccactatggatggcctcctgcgaggcttgcaag 471
            || ||| | ||     || || || ||  |||||||||     || || ||  |  || ||
Qy    292 GTCTGCAGTGATTACGCATCGGATATCAGTATCGAGTCTGGTCGTGTGAAGGATGTAAC 351

Db    472 gcgttctttaagaggaccattcaggaaaatatgaatacagctgccctgccaccaacgac 531
          || || || || |||| |||||||||| | |  || || |  || || || || ||  ||
Qy    352 GCCTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCAG 411

Db    532 tgtgagatcaccaacggaggcgcaagtcctgtcaggcctgccggttcatgaaatgcctc 591
          |||   |||   |||    ||||||||  ||| |||||||||||  |    ||| ||   |
Qy    412 TGTACAATCGATAAAAACGGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTAC 471

Db    592 aaagtggggatgctgaag 609
           ||||||| ||| |||||
Qy    472 GAAGTGGGAATGGTGAAG 489


RESULT   11
ID   Q10572 standard; DNA; 1047 BP.
AC   Q10572;
DT   09-APR-1991  (first entry)
DE   Human Natriuretic Peptide Receptor B.
KW   NPRB; ANP; BNP; CNP; kidney failure; heart failure; protein kinase;
KW   hyperaldosteronism; glaucoma; guanyl cyclase.
OS   Homo sapiens.
FH   Key               Location/Qualifiers
FT   Peptide           1..22
FT   /label= signal sequence
FT   Protein           12
FT   /label= mature NPBR
FT   Domain            23..455
FT   /label= extracellular domain
FT   /note= "binds natriuretic peptides A,B and C]"
FT   Domain            456..456
FT   /label= transmembrane domain
FT   Domain            479..1047
FT   /label= cytoplasmic domain
FT   /note= "GC and protien kinase activity"
FT   Modified -site  24..26
FT   /label= N-glycos_site
FT   Modified -site  35..37
FT   /label= N-glycos_site
FT   Modified -site  161..163
FT   /label= N-glycos_site
FT   Modified -site  195..197
FT   /label= N-glycos_site
FT   Modified -site  244..246
FT   /label= N-glycos_site
FT   Modified -site  277..279
FT   /label= N-glycos_site
FT   Modified -site  349..351
FT   /label= N-glycos_site
FT   Modified -site  600..602
FT   /label= N-glycos_site
PN   WO9100292-A.
PD   10-JAN-1991.
PF   22-JUN-1990; U03586.
PR   23-JUN-1989; US-370673
PA   (GETH ) GENENTECH INC.
PI   Chang M, Goeddel D, Lowe D;
DR   WPI; 91-036711/05.
DR   N-PSDB; Q10324.
PT   Natriuretic protein receptor B - for diagnosis and treatment of
PT   kidney failure, heart failure, hyperaldosteronism, glaucoma etc.
PS   Claim 3; Fig 1; 49pp; English.
CC   The sequence was derived from the DNA encoding natriuretic peptide
CC   receptor B, NPRB, having guanyl cyclase (GC) activity and protein
CC   kinase activity.  The DNA can be inserted into expression vectors
CC   for the prodn. of the protein, opt. after being mutated to produce
CC   NPRB analogues. The protein has a mol wt. of 115 kD (calculated Mr-
CC   114,952).  The protein (or variants) can be used in treatment of
CC   natriuretic peptide disorders, and also to isolate peptides using
CC   affinity chromatography.  Antibodies with affinity for NPRB can
CC   also be prepd
SQ   Sequence   1047 BP;    87 A;    15 C;    83 G;    51 T;

  Query Match             5.3%;  Score 66;  DB 2;  Length 1047;
  Best Local Similarity   9.2%;  Pred. No. 2.80e 27;
  Matches    93;  Conservative  268;  Mismatches 646;  Indels    9;  Gaps    9;

Db       18 rnnqarnntnavvnnhnnyawawnrvqnavanavnanfrannvdnrnvssnnngaesn 77
               :   |  :   |  ::     :     :          :   |:|  |:              :  :  ::
Cp     1068 ACTGGAATTGAGCAGGATCATGGCCTTGACACAGAGATATTCTTTGTGTTGGAGTTTTAA 1009

Db       78 ynannsavdnknyhdndnnnqngcvynaasvarnashwrnnnntaqavasqnsakndhyr 137
               :     :|::    :  :  :        ||.     |..|     ....:      |   |.|    :  .    .
Cp     1008 CTCTCGAAACCTTGAAGTAGTTGCCAGGAGCATGTCAAAGATTTCCAGAATTCCTTCTAC 949

Db      138 tnvrtqnsanknqnnvvtnhghnnwtnraannyndartddrnhyntnnqvnnanngsnns 197
                 :  |    :    :       ::|  :   :    :  |  ||   :  :  :  :  :  :    |    |:         |:
Cp      948 GCATTTCCCCTCATCCCTGTCCAGA-ACAAGATCTGGAGCAAAGATGAGCTTGCCGGGGT 890

Db      198 vnhnvyqrnnqqnnnathnnranqrnvynngnnnmnhnnnnnanrnnntngdyvnnyndv 257
               :  :   :            |        |   :    :              ::  |        :   :           |             |  ::    :  ::
Cp      889 GGTCAATTGAGCGCCACATCAGCCCCAT-CATTAACACCTCCATCCAACAGCTCTCCAAG 831

Db      258 nqnsnraqntratqrnwndnr-trnnananrnanntvnvntyrnnnnnnynnnnnrnnnr 316
               |  :    |    |:  |   :   :   :   :   |:                :  |     :   :       :            :          :      :
Cp      830 AGCCGCACTTGGTCGAACAGGCTGAGCTCCACAAAGCCGGGAATCTTCTTGGCCCAGCTG 771

Db      317 arndnqvnnqnsnmnnnaqcnydnnnnyanvnnntnnnqqtrndqnrnvnkmnqrryhqv 376
               |    :   |           :  :                :  |         |           |      ||       :            :   :   |:  ::  :
Cp      770 ATCATCTCTACCAACTCCTTCTCCCCAACTTCCTCAGGACATCATCATGGAGGCCTCG 711

Db      377 tqnvvmdknndrntdnvnwnmqdndsqdnnnaahysqanknnwwtqrnnnwvkqannsdn 435
                   ::  ::        :  |:  :              :   :     :              |:   :|     :     :     |              ::||     ::
Cp      710 GTGAAGGGGGTACTCCCCCGGCTCATCAGCACATCGGCGGCTCAGCTTTCAGGAGGGTG 651

Db      437 nncandnddnsed ktnnstnanvanqtqntnnmnqvssnnnnrknmnnknnasmnwrnr 495
                 ||   :    :     :|  :    :    :|   |   :     ||          .        .  .  .        .      .      .      .      .
Cp      650 AGCACTAGCTGCTCGGGGCTCAGGGCGTCCAGCAGCAGCTCCCGCACTCGGGGCGCGTGG 591

Db      496 wnnnnnnnsnryhknrqsrntnsnnqssyqsnmtahqkynnnentghnkqnvvankhvnk 555
                                           :::|    :         :    :     :    |      :       :   ::                  |:         :          :::
Cp      590 CCGCCACTTCTCTTGGCCTTGCCGGCACACTGCACCTCCTCGTCCCCACTTCTCTGTCTC 531

Db      556 krnnntrnvnnnnkhnndvnnnhntrnnqanndnnnnnvtnyenrgsnndnnnndsnnn 615
                  :          :   :         :: :    ::      :   |             |   :           |   :      :      :      :      :
Cp      530 CGCACAAGGCGGTACCCACATCTCTCTCTCCGGGAGCCACACTTCACCATTCCCACTTCG 471

Db      616 dwnnnrysnnndnvkqnannhnsnnsshqsnkssnnvvdsrnvnkntdyqnasnrstannd 675
              :::  ::          :  ::  :          :    :         :::|:       ::    |      :::   :    :    |::          :    :    |
```

```
Cp      470 TAACACTTCCGAAGTC-GGCAGG-CCTGGCAGCTCTTGCGCCGGTTTTTATCGATTGTAC 413

Db      676 dnnnnyskknntannnnsqnnnttqnnaadvysnqnnnnnannsqnnynnnndnsnkn 735
               :       :  ::               :             |            |   ::::                                        :            :          :
Cp      412 ACTGATTTGTAGCTGGACAAATATAATCATTATGTCCTTGAATGCTTCTTTTAAAAAAGG 353

Db      736 nvnkvrnqnrnynnnsndrtnnnnnnvnnmnrnwandnannndnqnnkqnnrnnknaqt 795
               :   ::          :   :       ::                  :              :     |  :                 :  |     :               :       :       |
Cp      352 CCTTACATCCTTCATACGACCAGATCCTATAGTGATATCCCGATCCGTAATCGCTCCACA 293

Db      796 snndnnnrmnnyannnnknvnnrtnaynnnkrkanannynnnnhsvannnkrqntvnan 855
               :                ::        |             :            |                :::|            :               :::               :   |       :
Cp      292 CAGCGCAGAAGTGAGTATCCCCTTTCAACCTGGACCAGTAAC-AGGGCTGCCCCAACGG 234

Db      856 andsvtnynsdnvqntansansnnmnvvtnnndnytnndanndnndvykvntnqdaymvv 915
                  :::    :     :      :                        :              :|         :   :|          |        :       ::::       |      :|   :::
Cp      233 -TTCCCACTAACCTTCCTTTTCAGTGTCTCTCTGTTTACAGGTAAGGTGTGTTCTAGCGA 175

Db      916 sqnnrnqnrhannnarmananndavssnnnnn-nnhdnnrnnnqvhtqnvnnnvqnkm 974
                          |           . . |            .    |          .        .        .        .        .      .  .          .      .      |                       :|    |::|    :
Cp      174 TCTTGCTTCACACCAGGGACTCTTTTGAGGTTCCGCATACAGATGTGATAACTGGCGATG 115

Db      975 nrynnngdtvntasnnnsngnanknhvssttkdandnngnnnnnrqdvnmkqkqk 1030
                  ::|      :    :     |. . .    .                            .   .||::       :             |                  |::       :|:|:
Cp      114 GACCACTAAAGGAGAAAGGTGCCCAGGTGTTGGCCACAACACATTGGGCTTGTGG 59


RESULT   12
ID   Q76269 standard; DNA; 6244 BP.
AC   Q76269;
DT   17-JUL-1995  (first entry)
DE   PhCMV*-1-controlled progesterone receptor
KW   tTA; transactivator, tetracycline controllable transactivator;
KW   conditional inactivation; homologous recombination, gene expression;
KW   gene regulation; gene therapy; tetracycline-resistance, tetR;
KW   transgenic animal; PhCMV*-1; promoter; tetO; CMV;
KW   progesterone receptor; ds.
OS   Human cytomegalovirus Towne.
PN   WO9429442-A.
PD   22 DEC-1994.
PF   14-JUN-1994; U06734.
PR   14-JUN-1993; US-076327.
PA   (BADI ) BASF AG.
PI   Bujard H,  Gossen M,  Salfeld JG,  Voss JW;
DR   WPI; 95-036472/05.
PT   Regulatory systems using tetracycline-controllable transactivator
PT   (tTA)  - useful for conditional inactivation or modulation of
PT   gene expression in a host cell or animal
PS   Disclosure; Page 54-58; 103pp; English.
CC   The minimal promoter PhCMV*-1 is based on the human CMV
CC   intermediate-early promoter (with enhancer region removed), fused
CC   to a tetracycline operator, tetO.  The promoter is used to
CC   express tTA transactivator in host cells.  In clone pUHD BGR3
CC   (given in Q76269), cDNA for rabbit progesterone receptor is under
CC   the control of PhCMV*-1.
SQ   Sequence   6244 BP;    1483 A;    1749 C;    1610 G;    1402 T;

  Query Match             5.0%;  Score 63;  DB 13;  Length 6244;
  Best Local Similarity  66.3%;  Pred. No. 3.03e-25;
  Matches   128;  Conservative    0;  Mismatches  65;  Indels    0;  Gaps    0;

Db     2365 tctgtgggatgaagcatcaggctgtcattatggtgtcctcacctgtgggactatagga 2424
              ||||   |  |||  |  |||||  ||  |  |||  |||||  |||      |  ||||     |  |||||||
Qy      293 TCTGCAGGGATTACGCATGGGATATCACTATCTACTTCTGGTCGTGTGAAGGATTTAAGG 352

Db     2425 tcttctttaaaggcaatggaaggcagcataactatttatatgctggaaggaatcact 2484
                |||  ||||||||      ||    ||||  ||    ||  |  |||  |  |||  |  |    |  ||||  |  |
Qy      353 CCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCAGT 412

Db     2485 gcattgttgataaaatccgcaggaaaaactgcccggcgtgtcgccttagaaagtactatc 2544
              |  |     |  |||||||  |||    |  ||  |  |||||  |||  ||  ||  |||  |  ||||    |
Qy      413 GTACAATCGATAAAAACCGGCCAAGAGCTGCCAGGCCTCCCGACTTCGGAAGTGTTACG 472
```

```
Db   2545 aagctggcatggt 2557
            ||  || |||||
Qy    473 AAGTGGGAATGGT 485


RESULT   13
ID   T06872 standard; DNA; 6244 BP.
AC   T06872;
DT   13-MAR-1996  (first entry)
DE   PhCMV*-1 promoter and rabbit progesterone receptor.
KW   Transactivator; tTA; tet operator; tetO; gene expression;
KW   tetracycline-responsive promoter; PhCMV*-1; human cytomegalovirus;
KW   transcription activation; eukaryotic cell; progesterone receptor;
KW   ds; ss.
OS   Chimeric synthetic;
OS   Chimeric Oryctolagus cuniculus.
PN   US5464758-A.
PD   07-NOV-1995.
PF   14-JUN-1993; 076726.
PR   14-JUN-1993; US-076726.
PA   (BUJA/) BUJARD H.
PA   (GOSS/) GOSSEN M.
PI   Bujard H,  Gossen M;
DR   WPI; 95-392612/50.
PT   Polynucleotide encoding trans-activator fusion protein contg. tet
PT   repressor - used to control expression of gene regulated by minimal
PT   promoter linked to tet operon, and vectors and cells where gene
PT   expression is regulated by tetracycline
PS   Disclosure; Fig 9A-C; 37pp; English.
CC   cDNA coding for the rabbit progesterone receptor was placed under
CC   control of the PhCMV*-1 promoter (see T06869), resulting in
CC   tetracycline-regulated expression vector pUHDpgr-3 contg. the
CC   sequence given in T06872. The construct was transfected into
CC   HeLa tTA cells expressing the tTA transactivator (T06867), to
CC   allow tetracycline-regulated expression of the receptor.
SQ   Sequence   6244 BP;   1483 A;   1750 C;   1609 G;   1402 T;

  Query Match              5.0%;  Score 63;  DB 17;  Length 6244;
  Best Local Similarity  66.3%;  Pred. No. 3.03e-25;
  Matches   128;  Conservative    0;  Mismatches  65;  Indels    0;  Gaps    0;

Db   2365 tctgtggggatgaagcatcaggctgtcattatggtgtcttccactgtggagctgtaaggg 2424
             |||  | ||  | ||||| || | ||| ||||| |||      | ||||    | |||||||
Qy    293 TCTGCAGGGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGGATGTAAGG 352

Db   2425 tttttttaaaagggggaatggaggggcagcataactatttatgtgctggaagaaatgact 2484
            ||| ||||||||     ||  |||| ||   || | ||| | ||| | |   | |||| | |
Qy    353 CCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTACAAATCAGT 412

Db   2485 gcattgttgataaaatccgcaggaaaaactgcccggcgtgtcgccttagaaagtgctgtc 2544
            | |    | ||||||| |||    | || | ||||| ||| || || ||| | ||||| |
Qy    413 GTACAATCGATAAAAACGGCGGCAAGAGCTGCCAGGCCTGCCGACTTCGGAAGTGTTACG 472

Db   2545 aagctggcatggt 2557
            ||  || || ||
Qy    473 AAGTGGGAATGGT 485


RESULT   14
ID   Q79934 standard; cDNA; 2992 BP.
AC   Q79934;
DT   01-AUG-1995  (first entry)
DE   Human liver hap cDNA.
KW   RAR-beta; retinoic acid receptor-beta; hap gene; hepatoma;
KW   hepatocellular carcinoma; hepatocyte; liver; hepatitis B virus;
KW   ss.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   misc_signal     119..121
FT   /*tag=  a
FT   /note= "untranslated start codon"
FT   misc_signal     126..128
FT   /*tag=  b
FT   /note= "untranslated start codon"
FT   misc_signal     178..180
FT   /*tag=  c
FT   /note= "untranslated start codon"
FT   misc_signal     237..239
FT   /*tag=  d
FT   /note= "untranslated start codon"
FT   CDS             322..1668
FT   /*tag=  e
FT   /product= RAR-beta
FT   polyA_signal    1954..2959
FT   /*tag=  f
PN   US5376530-A.
PD   27-DEC-1994.
PF   16-DEC-1987; 133687.
PR   16-DEC-1987; US-133687.
PR   17-DEC-1987; US-134130.
PR   20-JUN-1988; US-209009.
PR   30-NOV-1988; US-278136.
PR   30-MAR-1989; US-330405.
PR   21-AUG-1991; US-751612.
PR   30-MAR-1992; US-860577.
PR   11-DEC-1992; US-989902.
PR   22-JUL-1993; US-095706.
PA   (INSP ) INST PASTEUR.
PI   Brand N,  Chambon P,  De The HB,  Dejean A,  Krust A;
PI   Marchio A,  Petkovich M,  Tiollais P;
DR   WPI; 95-043461/06.
DR   P-PSDB; R66764.
PT   Detection of transformed hepatocyte(s) - by detecting a 2.5 kb
PT   mRNA species in hepatic tissue using DNA encoding a retinoic acid
PT   receptor-beta fragment.
PS   Disclosure; Fig.2A-2B; 35pp; English.
CC   A hepatitis B virus integration in a DNA fragment linked to
CC   hepatocellular carcinoma was used as a probe to clone cDNA from a
CC   human liver cDNA library.  The cellular gene, named hap, encoded a
CC   retinoic acid receptor identified as RAR-beta.
SQ   Sequence   2992 BP;   899 A;   652 C;   600 G;   841 T;

  Query Match              4.7%;  Score 59;  DB 13;  Length 2992;
  Best Local Similarity  64.8%;  Pred. No. 1.47e-22;
  Matches   129;  Conservative    0;  Mismatches  70;  Indels    0;  Gaps    0;

Db    561 ctgcttcgtctgcaggacaaatcatcagggtaccactatgggatcagcacctatgaaggg 620
            ||||    ||||||     ||   |    |||| || || ||||||||| ||| |   | ||||| ||
Qy    285 CTGCACTGTCTGCAGGGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGG 344

Db    621 atgtaagggctttttccgcaggaagtattcaggaagaatatgatttacacttgtcccaga 680
            ||||||||| ||||||     ||||| |||||       |||      ||| | |||||     |
Qy    345 ATGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTAC 404

Db    681 taggaactgtgttattaataaagtcaccaggaatcgatgccaatactgtcgactccagaa 740
             ||   | |||    ||   |||||  |      | ||   | |||||    ||| ||||| | |||
Qy    405 AAATCAGTGTACAATCGATAAAAACGGCGGCAAGAGCTGCCAGGCCTGCCGACTTCGGAA 464

Db    741 gtgctttgaagtgggaatg 759
            ||| |   ||.|||||||||
Qy    465 GTGTTACGAAGTGGGAATG 483


RESULT   15
ID   Q65572 standard; cDNA; 2992 BP.
AC   Q65572;
DT   11-JAN-1995  (first entry)
DE   Human liver hap cDNA.
KW   liver; hap; retinoic acid receptor; steroid; thyroid; hormone;
KW   hepatoma; retinoid; antibody; ss.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   CDS             322..1668
FT   /*tag=  a
FT   /product= RAR-beta retinoic acid receptor.
FT   misc_difference 916..918
FT   /*tag= b
FT   /transl_except= GCT encodes Glycine.
FT   misc_difference 1372..1374
FT   /*tag= c
FT   /transl_except= GGA encodes Alanine.
PN   US5317090-A.
PD   31-MAY-1994.
PF   16-DEC-1987; 133687.
PR   16-DEC-1987; US-133687.
PR   17-DEC-1987; US-134130.
PR   20-JUN-1988; US-209009.
PR   30-NOV-1988; US-278136.
PR   30-MAR-1989; US-330405.
PR   21 AUG-1991; US-751612.
PR   30-MAR-1992; US-860577.
PR   11-DEC-1992; US-989902.
PA   (INSP ) INST PASTEUR.
PI   Brand N,  Chambon P,  De THE HB,  Dejean A,  Krust A;
PI   Marchio A, Petkovich M,  Tiollais P;
DR   WPI; 94-176333/21.
PT   Antibody specific for retinoic acid receptor beta - useful for
PT   detecting, quantifying and identifying agonists and antagonists
PT   of retinoid activity.
PS   Disclosure; Figure 2; 35pp; English.
CC   The hap cDNA encodes a retinoic acid receptor called RAR-beta.  The
CC   hap gene is transcribed at low level in most human tissues, but the
CC   gene is overexpressed in prosate and kidney.  Six out of seven
CC   hepatoma or hepatoma-derived cell lines express a small hap
CC   transcript which is undetectable in normal adult and foetal livers
CC   but present in all non-hepatic tissues tested.
SQ   Sequence   2992 BP;    899 A;    653 C;    599 G;    841 T;

  Query Match             4.7%;  Score 59;  DB 11;  Length 2992;
  Best Local Similarity  64.8%;  Pred. No. 1.47e-22;
  Matches  129;  Conservative    0;  Mismatches  70;  Indels    0;  Gaps    0;

Db     561 ctgcttcgtctgccaggacaaatcatcagggtaccactatggggtcagcgcctgtgaggg 620
           ||||   ||||||   ||  |  |||| || || |||||||| ||| |  | ||||| ||
Qy     285 CTGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAGTCTGGTCGTGTGAAGG 344

Db     621 atgtaagggcttttccgcagaagtattcagaagaatatgatttacacttgtcacccaga 680
           |||||||| ||||||    ||||| |||||     |||    ||| | |||||   |
Qy     345 ATGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAATGATTATATTTGTCCAGCTAC 404

Db     681 taagaactgtgttattaataaagtcaccaggaatcgatgccaatactgtcgactccagaa 740
            ||  | |||   ||  |||||  |    | ||  | |||||   ||| ||||| | |||
Qy     405 AAATCAGTGTACAATCGATAAAAACCGGCGCAAGAGCTGCCAGGCCTGCCGACTTCGGAA 464

Db     741 gtgctttgaagtgggaatg 759
           ||| |  ||||||||||||
Qy     465 GTGTTACGAAGTGGGAATG 483

Search completed: Thu Dec 18 19:48:47 1997
Job time : 147 secs.
```

```
****************************************************************************

   ___    ___    _______    _______    _______    _______    _     _
  | |\   /  /|  |  ___  |  |  ____|   |  ___  |  |  _____|  | |   | |
  | | \ /  / |  | |   | |  | |        | |   | |  | |        | |   | |
  | |  \__/  |  | |___| |  | |____    | |___| |  | |        | |___| |
  | |        |  |  _____|  |____  |   |  __  _|  | |        |  ___  |
  | |        |  | |             | |   | |  \ \   | |        | |   | |
  | |        |  | |         ____| |   | |   \ \  | |_____   | |   | |
  |_|        |_|  |_|       |_______|  |_|    \_\ |_______|  |_|   |_|  (TM)

****************************************************************************

       Release 2.1D John F  Collins, Biocomputing Research Unit.
       Copyright (c) 1993  1994  1995  University of Edinburgh  U K.
                 Distribution rights by IntelliGenetics, Inc.

MPsrch_pp   protein - protein database search, using Smith-Waterman algorithm

Run on:          Thu Dec 18 09:03:07 1997;  MasPar time 10.31 Seconds
                                            498.420 Million cell updates/sec
Tabular output not generated.

Title:           >US-08-826-361A-21
Description:     (1-418) from US08826361A.pep
Perfect Score:   3116
Sequence:        1 MNYSIPSNVTNLEGGPGPQT.....   SMPLANLLMLLSHVPHARSA 418

Scoring table:   PAM 150
                 Gap 11

Searched:        101610 seqs, 12294212 residues

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database:        a-geneseq28
                   1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
                   8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
                   14:part14 15:part15 16:part16 17:part17 18:part18
                   19:part19 20:part20 21:part21

Statistics:      Mean 35.291;  Variance 145.647;  scale 0.242

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                SUMMARIES
                 %
Result          Query
  No.   Score   Match  Length DB  ID            Description              Pred. No
-------------------------------------------------------------------------------
    1    3099   99.5     485 21  W14724        Human oestrogen recep    0.00e+00
    2    2839   91.1     485 21  W14723        Rat oestrogen recepto    3.96e-277
    3    2920   90.5     485 21  W14725        Mouse oestrogen recep    2.96e-275
    4    1451   46.6     595  2  P70543        Human oestrogen recep    2.28e-133
    5    1422   45.6     334  2  P80925        Sequence of the human    2.21e-130
    6     736   23.6     433  2  P80931        Sequence of human est    2.41e-60
    7     725   23.3     521  2  P80930        Sequence of human est    3.09e-59
    8     614   19.7     777  2  P80919        Sequence of the prima    4.26e-49
    9     563   18.1      77 10  R53289        Human oestrogen recep    5.17e-43
   10     559   17.9     902  3  R12224        Rat androgen receptor    1.29e-42
   11     556   17.8     902  1  P91006        Rat androgen receptor    2.57e-42
   12     556   17.8     902  1  P93110        Rat androgen receptor    2.57e-42
   13     556   17.8     918  3  R12223        Human androgen recept    2.57e-42
   14     553   17.7     919  1  P90996        Human androgen recept    5.10e-42
   15     553   17.7     919 21  W14783        Androgen receptor        5.10e-42
   16     552   17.7     919  1  P93109        Human androgen recept    6.41e-42
   17     546   17.5      72  1  P91994        Part of oestrogen rec    2.53e-41
   18     545   17.5     630  3  R12230        TrpE/androgen recepto    3.18e-41
   19     536   17.2      69 13  R71566        Partial human estroge    2.48e-40
   20     521   16.7     984  2  P80927        Sequence of the human    7.60e-39
   21     505   16.2     446 13  R72484        H-2RIIBP.                2.91e-37
   22     505   16.2     446  2  R11837        H2RIIBP nuclear hormo    2.91e-37
   23     506   16.2     533  9  R39468        hRXR-beta1.              2.31e-37
   24     506   16.2     533 13  R72483        Human H-2RIIBP.          2.31e-37
   25     501   16.1     448  9  R39467        mRXR-beta.               7.22e-37
   26     501   16.1     510  8  R39469        hRXR-beta2               7.22e-37
   27     494   15.9     467  8  R39470        mRXR-alpha.              3.54e-36
   28     494   15.9     467  3  R13799        Mouse RXR-alpha retin    3.54e-36
   29     493   15.8     462  3  R13798        Human RXR-alpha retin    4.44e-36
   30     471   15.1     598  9  R48631        Sequence of nuclear r    6.52e-34
   31     464   14.9     463  8  R39471        mRXR-gamma.              3.18e-33
   32     461   14.8     463  3  R13800        Mouse RXR-gamma retin    6.27e-33
   33     449   14.4     463 20  W00605        Human retinoid X rece    9.42e-32
   34     438   14.1     628 16  R92057        Apoptopic cerebral ne    1.13e-30
   35     420   13.5     447 20  W03326        LXR-alpha, orphan mem    6.44e-29
   36     411   13.2     440  7  R33744        XR2.                     4.85e-28
   37     411   13.2     447  4  R20466        Mouse retinoic acid r    4.85e-28
   38     412   13.2     448 12  R66764        Human RAR-beta.          3.88e-28
   39     412   13.2     448 17  R85315        Human hepatoma retino    3.88e-28
   40     410   13.2     449 10  R55128        Retinoic acid recepto    6.07e-28
   41     411   13.2     454  2  R10182        Recombinant human gam    4.85e-28
   42     411   13.2     454  2  R10550        Human Retinoic Acid R    4.85e-28
   43     411   13.2     454 16  R84723        Human retinoic acid r    4.85e-28
   44     411   13.2     458  4  R20465        Mouse retinoic acid r    4.85e-28
   45     408   13.1     458  2  R10549        Skin-specific murine     9.50e-28

                                ALIGNMENTS

RESULT     1
ID   W14724 standard; Protein; 485 AA.
AC   W14724,
DT   08-JUN-1997  (first entry)
DE   Human oestrogen receptor beta (ER-beta).
KW   Orphan receptor; oestrogen receptor beta; ER-beta;
KW   nuclear receptor; prostate cancer, benign prostatic hyperplasia,
KW   ovary cancer, cardiovascular disease, osteoporosis;
KW   environment, pollutant.
OS   Homo sapiens.
FH   Key                 Location/Qualifiers
FT   Domain              104..169
FT   /label= DNA-binding_domain
FT   Domain              260..457
FT   /label= Ligand-binding_domain
PN   WO9709348-A2.
PD   13-MAR-1997.
PF   09-SEP-1996; E03933.
PR   08-SEP-1995; GB-018272.
PR   15-MAR-1996; GB-005550.
PR   11-APR-1996; GB-007532.
PR   08-MAY-1996, GB-009576.
PA   (KARO-) KARO BIO AB.
PI   Enmark E,  Gustafsson J,  Kuiper GG;
DR   WPI; 97-192842/17.
DR   N-PSDB, T62843.
PT   New isolated oestrogen receptor beta - used to develop prods. for
PT   treating e.g. cancers, CNS diseases, osteoporosis or cardiovascular
PT   disease
PS   Claim 1, Fig 13A, 45pp; English.
CC   A novel human orphan receptor (W14724) is related to the known
CC   oestrogen receptor ER-alpha, and has been designated ER-beta.
CC   It is an oestrogen receptor-related nuclear receptor.  The ER-beta
CC   amino acid sequence was deduced from a cDNA clone (T62843) isolated
CC   from a human testis cDNA library.  Rat, human and mouse ER beta
CC   (W14723-25) can be used to isolate molecules for use in the
CC   treatment of cardiovascular diseases, central nervous system
CC   diseases, osteoporosis, prostate or ovarian cancer or benign
CC   prostatic hyperplasia and to test environmental chemicals for
CC   oestrogenic activity.
SQ   Sequence   485 AA;

  Query Match              99.5%;  Score 3099;  DB 21;  Length 485;
  Best Local Similarity 100.0%;  Pred. No. 0.00e+00;
  Matches  415;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Db        9 mnysipsnvtnleggpgrqttspnvlwptpghlsplvvhrqlshlyaepqkspwcearsl 68
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        1 MNYSIPSNVTNLEGGPGRQTTSPNVLWPTPGHLSPLVVHRQLSHLYAEPQKSPWCEARSL 60

Db       69 ehtlpvnretlkrkvsgnrcaspvtgpgskrdahfcavcsdyasgyhygvwscegckaff 128
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       61 EHTLPVNRETLKRKVSGNRCASPVTGPGSKRDAHFCAVCSDYASGYHYGVWSCEGCKAFF 120

Db      129 krsiqghndyicpatnqctidknrrkscqacrlrkcyevgmvkcgsrrercgyrlvrrqr 188
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      121 KRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQR 180

Db      189 sadeqlhcagkakrsgghaprvrellldalspeqlvltlleaepphvlisrpsapfteas 248
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      181 SADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQLVLTLLEAEPPHVLISRPSAPFTEAS 240

Db      249 mmmsltkladkelvhmiswakkipgfvelslfdqvrllescwmevlmmglmwrsidhpgk 308
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      241 MMMSLTKLADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGK 300

Db      309 lifapdlvldrdegkcvegileifdmllattsrfrelklqhkeylcvkamillnssmypl 368
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      301 LIFAPDLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPL 360

Db      369 vtatqdadssrklahllnavtdalvwviaksgissqqqsmrlanllmllshvrha 423
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      361 VTATQDADSSRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHA 415

RESULT    2
ID   W14723 standard; Protein; 485 AA.
AC   W14723;
DT   08-JUN-1997  (first entry)
DE   Rat oestrogen receptor beta (ER-beta).
KW   Orphan receptor; oestrogen receptor beta, ER-beta;
KW   nuclear receptor; prostate cancer; benign prostatic hyperplasia;
KW   ovary cancer; cardiovascular disease, osteoporosis;
KW   environment; pollutant.
OS   Rattus sp.
FH   Key              Location/Qualifiers
FT   Domain           104..169
FT   /label= DNA-binding_domain
FT   Domain           259..457
FT   /label= Ligand-binding_domain
PN   WO9709348-A2.
PD   13-MAR-1997.
PF   09-SEP-1996; E03933.
PR   08-SEP-1995; GB-018272.
PR   15-MAR-1996; GB-005550.
PR   11-APR-1996; GB-007532.
PR   08-MAY-1996; GB-009576.
PA   (KARO-) KARO BIO AB.
PI   Enmark E,  Gustafsson J,  Kuiper GG;
DR   WPI; 97-192842/17.
DR   N-PSDB; T62842.
PT   New isolated oestrogen receptor beta - used to develop prods. for
PT   treating e.g. cancers, CNS diseases, osteoporosis or cardiovascular
PT   disease
PS   Claim 1; Fig 1; 45pp; English.
CC   A novel rat orphan receptor (W14723) is related to the known
CC   oestrogen receptor ER-alpha, and has been designated ER-beta.
CC   It is an oestrogen receptor-related nuclear receptor.  The ER-beta
CC   amino acid sequence was deduced from a cDNA clone (T62842) isolated
CC   from a rat prostate cDNA library.  Rat, human and mouse ER-beta
CC   (W14723-25) can be used to isolate molecules for use in the
CC   treatment of cardiovascular diseases, central nervous system
CC   diseases, osteoporosis, prostate or ovarian cancer or benign
CC   prostatic hyperplasia and to test environmental chemicals for
CC   oestrogenic activity.
SQ   Sequence   485 AA;
```

```
  Query Match            91.1%;  Score 2938;  DB 21;  Length 485;
  Best Local Similarity  89.1%;  Pred. No. 3.96e-277;
  Matches  369;  Conservative   29;  Mismatches   16;  Indels    0;  Gaps    0;

Db        9 mnysvpgstsnldggpvrlstspnvlwptsghlsplathcqssllyaepqkspwcearsl 68
            ||||:|:: :||:||| | :||||||||:||||||  | | | | ||||||||||||||||
Qy        1 MNYSIPSNVTNLEGGPGRQTTSPNVLWPTPGHLSPLVVHRQLSHLYAEPQKSPWCEARSL 60

Db       69 ehtlpvnretlkrklsgsslrmpvtspnakrdahfcpvcsdyasgyhygvwscegckaff 128
            ||||||||||||||:||: '||||||:|  :|||||||:|||||||||||||||||||||
Qy       61 EHTLPVNRETLKRKVSGNRCASPVTGPGSKRDAHFCAVCSDYASGYHYGVWSCEGCKAFF 120

Db      129 krsiqghndyicpatnqctidknrrkscqacrlrkcyevgmvkcgsrrercgyrivrrqr 188
            ||||||||||||||||||||||||||||||||||||||||:|||||||||||||:|||||
Qy      121 KRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQR 180

Db      189 ssseqvhclskakrngghaprvkelllstlspeqlvltlleaeppnvlvsrpsmpfteas 248
            |: ||:|| :||||:|||||||:|||| :||||||||||||||||:||:|||| ||||||
Qy      181 SADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQLVLTLLEAEPPHVLISRPSAPFTEAS 240

Db      249 mmmsltkladkelvhmigwakkipgfvelslldqvrllescwmevlmvglmwrsidhpgk 308
            |||||||||||||||||:|||||||||||||:||||||||||||||||:||||||||||||
Qy      241 MMMSLTKLADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGK 300

Db      309 lifapdlvldrdegkcvegileifdmllattsrfrelklqhkeylcvkamillnssmypl 368
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      301 LIFAPDLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPL 360

Db      369 asanqeaessrklthllnavtdalvwviaksgissqqqsvrlanllmllshvrh 422
             :| |:|:|||||:|||||||||||||||||||||||||:||||||||||||||
Qy      361 VTATQDADSSRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRH 414

RESULT    3
ID   W14725 standard; Protein; 485 AA.
AC   W14725;
DT   08-JUN-1997  (first entry)
DE   Mouse oestrogen receptor beta (ER-beta).
KW   Orphan receptor, oestrogen receptor beta; ER-beta;
KW   nuclear receptor; prostate cancer; benign prostatic hyperplasia;
KW   ovary cancer; cardiovascular disease; osteoporosis;
KW   environment, pollutant.
OS   Mus sp.
FH   Key              Location/Qualifiers
FT   Domain           104..169
FT   /label= DNA-binding_domain
FT   Domain           260..457
FT   /label= Ligand-binding_domain
PN   WO9709348-A2.
PD   13-MAR-1997.
PF   09-SEP-1996; E03933.
PR   08-SEP-1995; GB-018272.
PR   15-MAR-1996; GB-005550.
PR   11-APR-1996; GB-007532.
PR   08-MAY-1996; GB-009576.
PA   (KARO-) KARO BIO AB.
PI   Enmark E,  Gustafsson J,  Kuiper GG;
DR   WPI; 97-192842/17.
DR   N-PSDB; T62844.
PT   New isolated oestrogen receptor beta - used to develop prods. for
PT   treating e.g. cancers, CNS diseases, osteoporosis or cardiovascular
PT   disease
PS   Claim 1; Fig 14A; 45pp; English.
CC   A novel mouse orphan receptor (W14725) is related to the known
CC   oestrogen receptor ER-alpha, and has been designated ER-beta.
CC   It is an oestrogen receptor-related nuclear receptor.  The ER beta
CC   amino acid sequence was deduced from a cDNA clone (T62844).
CC   Rat, human and mouse ER-beta (W14723-25) can be used to isolate
CC   molecules for use in the treatment of cardiovascular diseases,
CC   central nervous system diseases, osteoporosis, prostate or ovarian
CC   cancer or benign prostatic hyperplasia and to test environmental
CC   chemicals for oestrogenic activity
SQ   Sequence   485 AA;
```

```
  Query Match             90.5%;  Score 2820;  DB 21;  Length 485;
  Best Local Similarity   88.6%;  Pred. No. 2.96e-275;
  Matches  367;  Conservative   31;  Mismatches   16;  Indels    0;  Gaps    0;

Db        9 mnysvpsstgnleggpvrqtaspnvlwptsghlsplathqgsslly aepqkspwcearsl 68
            ||||:||:  |||||| |||:||||||||:||||||  |   | ||||||||||||||||
Qy        1 MNYSIPSNVTNLEGGPGRQTTSPNVLWPTPGHLSPLVVHRQLSHLYAEPQKSPWCEARSL 60

Db       69 ehtlpvnretlkrklggsgcaspvtspstkrdahfcavcsdyasgyhygvwscegckaff 128
            ||||||||||||||::|: |||||:|::|||||||||||||||||||||||||||||||
Qy       61 EHTLPVNRETLKRKVSGNRCASPVTGPGSKRDAHFCAVCSDYASGYHYGVWSCEGCKAFF 120

Db      129 krsiqghndyicpatnqctidknrrksqqacrlrkcyevgmvkcgsrrercgyrlvrrqr 188
            |||||||||||||||||||||||||||||||:||||||||||||||||||||||:|||||
Qy      121 KRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQR 180

Db      189 sasegvhclnkakrtsghtprvkelllnslspeqlvltlleaeppnvlvsrpsmpftqas 248
            || ||:||  ||||::||:|||:||||::|||||||||||||||||:||:|||| ||||||
Qy      181 SADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQLVLTLLEAEPPHVLISRPSAPFTEAS 240

Db      249 mmmsltkladkelvhmigwakkipgfvelslldqvrllescwmevlmvglmwrsidhpgk 308
            |||||||||||||||||:|||||||||||||:|||||||||||||||:||||||||||||
Qy      241 MMMSLTKLADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGK 300

Db      309 lifapdlvldrdegkcvegileifdmllattarfrelklqhkeylcvkamillnssmyhl 368
            |||||||||||||||||||||||||||||||:|||||||||||||||||||||||||| |
Qy      301 LIFAPDLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPL 360

Db      369 atasqeaessrklthllnavtdalvwviskrsissqqqsvrlanllmllshvrh 422
             ||:|:|:|||||:||||||||||||||:|| ||||||:||||||||||||||
Qy      361 VTATQDADSSRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRH 414


RESULT    4
ID   P70543 standard; protein; 595 AA.
AC   P70543;
DT   24-FEB-1991  (first entry)
DE   Human oestrogen receptor protein.
KW   Human oestrogen receptor protein, expression system, assay.
OS   Homo sapiens.
PN   WO8705049-A.
PD   27-AUG-1987.
PF   18-FEB-1987; U00341.
PR   20-FEB-1986; US-833829.
PA   (CALB-) California Biotechnology Inc.
PI   Shine J;
DR   WPI; 87-250213/35
DR   N-PSDB; N70880
PT   Expression system for vertebrate steroid receptor protein - comprising
PT   DNA sequence encoding the protein linked to control sequences in
PT   eucaryotic hosts
PS   Disclosure; Fig. 1-1 - 1-2; 26pp; English.
CC   steroid receptor protein
CC   The protein is expressed in an expression system from cDNA operably
CC   linked to control sequences compatible with eukaryotic host cells
CC   This method allows expression under conditions which favour appropriate
CC   post-translational processing.  It also produces large amts. of purified
CC   protein useful in the design of agonist and antagonist cpds. for study
CC   of the mechanism of action of the steroid binding proteins in general
CC   and for use in diagnostic assays for the proteins or antibodies to
CC   them.  These assays are important in, eg the diagnosis of tumour
CC   sensitivities to steroid metabolism.  Suitable host cells are VERO,
CC   HeLa and CHO cells.
SQ   Sequence   595 AA;

  Query Match             46.6%;  Score 1451;  DB 2;  Length 595;
  Best Local Similarity   57.1%;  Pred. No. 2.28e-133;
  Matches  194;  Conservative   84;  Mismatches   47;  Indels   15;  Gaps   11;

Db      178 sakptrycavcndyasgyhygvwscegckaffkrsiqghndymcpatnqctidknrrksc 237
            | :   :||||:||||||||||||||||||||||||||||||:|||||||||||||||||
Qy       89 SKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSC 148

Db      238 qacrlrkcyevgmmkggirkdrrggrmlkhkrqrddgegrgevgsagdmraanlwpsplm 297
            |||||||||||||.| | |..| | |...  |||   | .   :: : :|::     :|
Qy      149 QACRLRKCYEVGMVKCGSRRERCGYRLVR--RQRSADE-QLHCAGKA-KRSGG--HAP-R 201

Db      298 ikrskknslalsltadqmvsalldaepp-ilyseydptrpfseasmmglltnladrelvh 356
            : |      |   :|:::|:| :||:||||  :| |     |: ||:||||  ||:|||:|:||
Qy      202 V-RELL--LD-ALSPEQLVLTLLEAEPPHVLISR--PSAPFTEASMMMSLTKLADKELVH 255

Db      357 minwakrvpgfvdltlhdqvhllecawleilmiglvwrsmehpvkllfapnllldrnqgk 416
            ||.|||..||||.|.| |||.|||  |.|.||.||.||:.|| ||.|||:|:|||:::||
Qy      256 MISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGK 315

Db      417 cvegmveifdmllatssrfrmmnlqgeefvclksiillnsgvytflsstlksleekdhih 476
            ||||::|||||||||:|||| ::||  |::|:|::|||||::| ::::|  : :     :
Qy      316 CVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSRKLA 374

Db      477 rvldkitdtlihlmakagltlqqqhqrlaqlllilshirh 516
             ..|. .||.|: :.||. .. |||  ||| ||..|||:||
Qy      375 HLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRH 414


RESULT    5
ID   P80925 standard; protein; 334 AA.
AC   P80925;
DT   29-DEC-1990  (first entry)
DE   Sequence of the human oestrogen receptor
KW   Hormone receptor; hormone-binding, transcription activation;.
OS   Homo sapiens.
PN   WO8803168-A.
PD   05-MAY-1988.
PF   23-OCT-1987; U02782.
PR   20-OCT-1987; US-108471.
PA   (SALK) Salk Inst for Biol Stud.
PI   Evans RM, Weinberger CA, Hollenberg SM, Giguere V;
DR   WPI; 88-133242/19.
PT   Recombinant DNA encoding hormone recpetors -
PT   comprising glucocorticoid, mineralocorticoid, thyroid hormone
PT   and novel hormone receptors
PS   Example, Fig III-2, 243pp, English.
CC   DNAs encoding hormone receptors and the hormone receptors themselves are
CC   claimed  The DNA can be used to make the hormone receptor proteins and
CC   functional modified forms in quantities not previously possible. The
CC   receptor proteins can be used to screen cpds. for receptor agonist or
CC   receptor antagonist activity. They can also be used in diagnostic assays.
SQ   Sequence   334 AA;

  Query Match             45.6%;  Score 1422;  DB 2;  Length 334;
  Best Local Similarity   57.5%;  Pred. No. 2.21e-130;
  Matches  192;  Conservative   80;  Mismatches   46;  Indels   16;  Gaps   12;

Db        1 ycavcndyasgyhygvwscegckaffkrsiqghndymcpatnqctidknrrkscqacrlr 59
            :||||:||||||||||||||||||||||||||||||:|||||||||||||||||||||
Qy       95 HCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLR 154

Db       60 kcyrvgmmkggirkdrrggrmlkhkrqrddgegrgevgsagdmraanlwpsplmikrskk 119
            ||| |||:| | |..| | |...  |||   | .   .: : :|::     :|  : |
Qy      155 KCYEVGMVKCGSRRERCGYRLVR--RQRSADE-QLHCAGKA-KRSGG--HAP-RV-RELL 206

Db      120 nslalsltadqmvsalldaepp-ilyseydptrpfseasmmglltnladrelvhminwak 178
              |   :|:::|:| :||:||||  :| |     |: ||:||||  ||:|||:||||:.||
Qy      207 --LD-ALSPEQLVLTLLEAEPPHVLISR--PSAPFTEASMMMSLTKLADKELVHMISWAK 261

Db      179 rvpgfvdltlhdqvhllecawleilmiglvwrsmehpvkllfapnllldrnqgkcvegmv 238
            ::|||||:|:| |||:|||  |:|:||:||:|||::|| ||:||||:|:||||::|||||::
Qy      262 KIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGIL 321

Db      239 eifdmllatssrfrmmnlqgeefvclksiillnsgvytflsstlksleekdhihrvldki 298
            |||||||||.||||  ..||  |.:|:|..|||||..| ::::|  : :     : ::|: :
Qy      322 EIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSRKLAHLLNAV 380

Db      299 tdtlihlmakagltlqqqhqrlaqlllilshirh 332
            ||.|. ..:.....  |||  ||| ||..|||:||
Qy     381 TDALVWVIAKSGISSQQQSMRLANLLMLLSHVRH 414


RESULT    6
ID   P80931 standard; protein; 433 AA.
AC   P80931;
DT   29-DEC-1990  (first entry)
DE   Sequence of human estrogen-related receptor protein (hERR2)
KW   Hormone receptor; hormone-binding; transcription activation.
OS   Homo sapiens
PN   WO8803168-A.
PD   05-MAY-1988.
PF   23-OCT-1987; U02782.
PR   20-OCT-1987; US-108471.
PA   (SALK) Salk Inst for Biol Stud.
PI   Evans RM, Weinberger CA, Hollenberg SM, Giguere V;
DR   WPI; 88-133242/19
DR   N-PSDB; N80922.
PT   Recombinant DNA encoding hormone recepetors -
PT   comprising glucocorticoid, mineralocorticoid, thyroid hormone
PT   and novel hormone receptors
PS   Claim 17; Fig V-2(B)-1 and -2; 243pp; English.
CC   DNAs encoding hormone receptors and the hormone receptors themselves are
CC   claimed. The DNA can be used to make the hormone receptor proteins and
CC   functional modified forms in quantities not previously possible. The
CC   receptor proteins can be used to screen cpds. for receptor-agonist or
CC   receptor-antagonist activity. They can also be used in diagnostic assays
CC   Also claimed is pure DNA encoding oestrogen-related receptors hERR1 and
CC   hERR2 and hERR1 and hERR2. The new hERR1 and hERR2 receptors will provide
CC   the basis for development of an assay system that will lead to the
CC   identification of novel hormones.
SQ   Sequence   433 AA;

  Query Match            23.6%;  Score 736;  DB 2;  Length 433;
  Best Local Similarity  43.6%;  Pred. No. 2.41e-60;
  Matches  116;  Conservative   55;  Mismatches  84;  Indels  11;  Gaps   9;

Db     101 rlclvcgdiasgyhygvasceackaffkrtiqgnieyscpatneceitkrrrksgqacrf 160
           ::| ||:| |||||||| |||:|||||||||:|||: :| |||||:| | | |||||||||:
Qy      94 HFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRL 153

Db     161 mkclkvgmlkeqvrldrv--r-ggrqky-krrldsesspylslqisppakkpltkivs-- 214
            ||  |||:| | | :|    |    ||:     :|        :  |    :|  :  |   :|
Qy     154 RKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPE 213

Db     215 yl-l-vaepdklyam-ppddvpegdikalttlcdladrelvfliswakhipgfsnltlgd 271
            | |  : |::      : :     |   :      : :|  |||:||| :||||| |||| :|:| |
Qy     214 QLVLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 273

Db     272 qmsllqsawmeilllgivyrslpyddklayaedyimdeehsrlvglleiyrailqlvrr 330
           |: ||:| |||:|::|:: ||:      || :| |  ::| : :: | |:||::  :|     |
Qy     274 QVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSR 333

Db     331 ykklkvekeefvmlkalalansdsmy 356
           :: ||::  |:: :||: | || |||
Qy     334 FRELKLQHKEYLCVKAMILLNSSMY 358


RESULT    7
ID   P80930 standard; protein; 521 AA.
AC   P80930;
DT   10-MAR-1993  (revised)
DT   29-DEC-1990  (first entry)
DE   Sequence of human estrogen related receptor protein (hERR1)
KW   Hormone receptor; hormone binding; transcription activation.
OS   Homo sapiens
PN   WO8803168-A.
PD   05-MAY-1988.
PF   23-OCT-1987; U02782.
PR   20-OCT-1987; US-108471.
PA   (SALK) Salk Inst for Biol Stud.
PI   Evans RM, Weinberger CA, Hollenberg SM, Giguere V;
DR   WPI; 88-133242/19
```

```
DR   N-PSDB; N80921.
PT   Recombinant DNA encoding hormone receptors -
PT   comprising glucocorticoid, mineralocorticoid, thyroid hormone
PT   and novel hormone receptors
PS   Claim 17; Fig V-1(B)-1 and -2; 243pp; English.
CC   DNAs encoding hormone receptors and the hormone receptors themselves are
CC   claimed. The DNA can be used to make the hormone receptor proteins and
CC   functional modified forms in quantities not previously possible. The
CC   receptor proteins can be used to screen cpds. for receptor-agonist or
CC   receptor-antagonist activity. They can also be used in diagnostic assays.
CC   Also claimed is pure DNA encoding oestrogen-related receptors hERR1 and
CC   hERR2 and hERR1 and hERR2. The new hERR1 and hERR2 receptors will provide
CC   the basis for development of an assay system that will lead to the
CC   identification of novel hormones.
SQ   Sequence   521 AA;

  Query Match            23.3%;  Score 725;  DB 2;  Length 521;
  Best Local Similarity  40.1%;  Pred. No. 3.09e-59;
  Matches  130;  Conservative   67;  Mismatches 117;  Indels  10;  Gaps   7;

Db     174 rlclvcgdvasgyhygvasceackaffkrtiqgsieyscpasneceitkrrrkacqacrf 233
           ::| ||:| |||||||| |||:|||||||||:|||    :| |||:|:| | | |||: ||||:
Qy      94 HFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRL 153

Db     234 tkclrvgmlkegvrldrvrggrqkykrrpevdplpfpgpfpagplavaggprktaap 291
            ||  |||:| | | :|  | |     ::|:  : |  :| :  ::|   :         | :
Qy     154 RKCYEVGMVKCGSRRERC-GYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSP 212

Db     292 nalvshllvvepeklyampdpagpdghlpavatlcdlfdrelvvtiswaksipgfsslsl 351
           : ||  ||  ||  :  :: |::|     : : :|  | |:|:|  ||||| |||'  |||
Qy     213 EQLVLTLLEAEPPHVL-ISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSL 271

Db     352 sdqmsvlqsvwmevlvlgvaqrslplqdelafaedlvld-eegaraaglgelgaallqlv 410
            ||: :|:| ||||||::|:  ||:       | || ||||| :||      |: |:   ||
Qy     272 FDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATT 331

Db     411 rrlqalrlereeyvllkalalansdsvhiedeprlwsscekllh  eal---leyeaqra 465
            |:: |:|:: ||: :||: | ||      .     :     |  || |   :|:    |    :::
Qy     332 SRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKS 391

Db     466 qpgqgaerrragrllltlpllrqt 489
           | ..   .  | . ||. |. .|::
Qy     392 GISSQQQSMRLANLLMLLSHVRHA 415


RESULT    8
ID   P80919 standard; protein; 777 AA.
AC   P80919;
DT   29-DEC-1990  (first entry)
DE   Sequence of the primary protein sequence of human glucocorticoid
DE   receptor (hGR)
KW   Hormone receptor; hormone-binding; transcription activation.
OS   Homo sapiens.
PN   WO8803168-A.
PD   05-MAY-1988.
PF   23-OCT-1987; U02782.
PR   20-OCT-1987; US-108471.
PA   (SALK) Salk Inst for Biol Stud.
PI   Evans RM, Weinberger CA, Hollenberg SM, Giguere V;
DR   WPI; 88-133242/19.
DR   N-PSDB; N80916.
PT   Recombinant DNA encoding hormone recepetors -
PT   comprising glucocorticoid, mineralocorticoid, thyroid hormone
PT   and novel hormone receptors
PS   Claim 24; Fig I-2(1) to (2); 243pp; English.
CC   DNAs encoding hormone receptors and the hormone receptors themselves are
CC   claimed. The DNA can be used to make the hormone receptor proteins and
CC   functional modified forms in quantities not previously possible. The
CC   receptor proteins can be used to screen cpds. for receptor-agonist or
CC   receptor-antagonist activity. They can also be used in diagnostic assays.
SQ   Sequence   777 AA;

  Query Match            19.7%;  Score 614;  DB 2;  Length 777;
  Best Local Similarity  33.4%;  Pred. No. 4.26e-48;
  Matches    97;  Conservative   84;  Mismatches   93;  Indels   16;  Gaps   15;

Db    397 smrpdvssppsssstattgp-p pkl-lvcsdeasgchygvltcgsckvffkravegqhn 454
          :::  ||:    :|: :  |:     : :| |||| ||  |||| :| :|| |||||:::|:::
Qy     70 TLFPKVSGNFCASPVIGPGSKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHND 129

Db    455 ylcagrnd-iidkirrknopa-ryrkclqagmnl-earktk-k-kikgiqqattg-v-sq 509
          |:|:: |:| ||| |||:| ||| ||| : ||     :|: :   ::   |:::    :
Qy    130 YICPATNQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCA 189

Db    510 otsenpgnkt-ivpatlpq-ltpt-lv-sllevieppvlyagydssvpdstwrimttlnm 565
            :  .|  .  |    | . |.| || .|||    | ||  .    .:  ...  .| .| .
Qy    190 GKAKRSGGHAPRVRELLLDALSPEQLVLTLLEAEPPHVLISRPSAPFTEAS--MMMSLTK 247

Db    566 lggrqvicavkwakaipgfrnlhlddqmtllqyswmflmafalgwrsyrqssanllcfap 625
          |: ::::  : ||| |||| :| | ||: ||.  ||  ..   .| |||   ::: :|: |||
Qy    248 LADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPG-KLI-FAP 305

Db    626 dlilne-qrmtlp-cmydqckhmlyvsselhrlqvsyeeylcmktlllss 674
          ||::: . . . .  . .     :|  .|  .. | .   ||||:|...||.|
Qy    306 DLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNS 355


RESULT    9
ID   R53289 standard; protein; 77 AA.
AC   R53289;
DT   30-DEC-1994  (first entry)
DE   Human oestrogen receptor DNA-binding domain (185-261).
KW   Human, oestrogen receptor; DNA-binding domain; hER;
KW   monoclonal antibody, diagnosis; antiserum; cancer; hapten;
KW   immunogen.
OS   Homo sapiens.
FH   Key              Location/Qualifiers
FT   Peptide          17..31
FT   /label= peptide_1
FT   /note= "hER 201-215"
FT   Peptide          46  61
FT   /label= peptide_2
FT   /note= "hER 231-245"
FT   Peptide          63..77
FT   /label= peptide_3
FT   /note= "hER 247-261"
PN   US5312752-A.
PD   17-MAY-1994.
PF   31-JUL-1989; 388091
PR   31-JUL-1989; US-388091.
PP   01-NOV-1991; US-784626
PA   (UYBO-) UNIV BOSTON.
PI   Traish AM,  Wotiz HH;
DR   WPI; 94-159123/19.
PT   Monoclonal antibody to human oestrogen receptor protein -
PT   specific for activated forms, useful for diagnosis, esp. in
PT   breast cancer assessment
PS   Disclosure; Fig 1; 41pp; English.
CC   The DNA-binding region of human oestrogen receptor (hER) is
CC   encompassed by the amino acids 180 251 in the native protein.
CC   Three peptides chosen as haptens and immunogens to obtain specific
CC   antisera, are indicated in the features table.
SQ   Sequence   77 AA,

  Query Match              18.1%;  Score 563;  DB 10;  Length 77;
  Best Local Similarity  89.6%;  Pred. No. 5.17e-43;
  Matches    69;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Db      1 cavcndyasgyhygvwscegckaffkrsiqghndymcpatnqctidknrrkscqacrlrk 60
          ||||:|| ||||||||||||||||||||||||||||:|||||||||||||||||||||||
Qy     96 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 155

Db     61 cyevgmmkggirkdrrq 77
          ||||||:| | |::| |
Qy    156 CYEVGMVKCGSPPERCG 172
```

```
RESULT   10
ID   R12224 standard; Protein; 902 AA.
AC   R12224,
DT   20-AUG-1991  (first entry)
DE   Rat androgen receptor.
KW   rAR; DNA-binding protein; steroid hormone.
OS   Rattus rattus.
FH   Key              Location/Qualifiers
FT   Domain           540..610
FT   /label= DNA-binding domain
FT   /note= "cysteine-rich"
PN   WO9107423-A.
PD   30-MAY-1991.
PF   19-OCT-1990; U06015.
PR   17-NOV-1989, US-438775.
PA   (ARCH-) ARCH DEV CORP.
PI   Liao S, Chang C;
DR   WPI; 91-178048/24.
DR   N-PSDB; Q11112.
PT   Androgen receptor and TR2 DNA binding proteins - DNA sequences
PT   and antibodies for detection and quantification methods
PS   Claim 25; Fig 3; 79pp; English.
CC   This sequence was deduced from a cDNA clone isolated by screening
CC   a rat ventral prostate lambda gt11 library in E.coli Y1090. The
CC   sequence is very similar to that of human AR and in
CC   the DNA-binding domain it is identical to that of hAR DNA-binding
CC   domain. Homology comparisons with other known steroid receptors
CC   indicate that rAR is more closely related to glucocorticoid,
CC   mineralo-corticoid and progesterone receptors than to v-erb-A or to
CC   receptors for oestrogen, vitamin D and thyroid hormones.
SQ   Sequence   902 AA;

  Query Match              17.9%;  Score 559;  DB 4;  Length 902;
  Best Local Similarity  33.5%;  Pred. No. 1.29e 42;
  Matches    88;  Conservative   72;  Mismatches   92;  Indels   11;  Gaps    8;

Db    542 clicgdeasgchygaltcgsckvffkraaegkqkylcasrndctidkfrrkncpscrlrk 601
          | .|.| ||| |||   :| :|| ||||: :|    |:|:: |:|||||| |||:| :||||
Qy     96 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 155

Db    602 cyeagmtl-garklkk---lgnlklqeegenssagsptedpsqkmtvshi--egyecgpl 655
          ||| ||    |.|.  .    .|          . .    ||  .     ..   |   .  ::   : :
Qy    156 CYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 215

Db    656 flnvleaiepgvvcaghdnnqpdsfaallsslnelgerqlvhvvkwakalpgfrnlhvdd 715
            | :|||  | |:   . . .   |  .  |:.. ||   |..::|||:: ||| :||| :| : |
Qy    216 VLTLLEAEPPHVL-ISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 273

Db    716 qmavieyswmglmvfamgwrsftnvnsrmlyfapdlvfn-eyrmhksrmysqcvrmrhls 774
          |    |  ||  :   .  |||          . ||||||. .  .    :         :    :
Qy    274 QVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDMLLATT 331

Db    775 qefgwlqitpqeflcmkalllfs 797
             |  |  .   |.||.||..|.:
Qy    332 SRFRELKLQHKEYLCVKAMILLN 354


RESULT   11
ID   P91006 standard; protein; 902 BP.
AC   P91006;
DT   28-FEB-1990  (first entry)
DE   Rat androgen receptor DNA clone.
KW   Androgen receptor, TR2 polypeptide;
OS   Rat.
FH   Key              Location/Qualifiers
FT   Region           1..902
FT   /*tag=  a
FT   /product=98 kD polypeptide
FT   Region           170..902
FT   /*tag=  b
FT   /product=79 kD polypeptide
PN   WO8909223-A.
```

```
PD   05-OCT-1989
PF   24-MAR-1989; U01238.
PR   30-MAR-1988; US-176107.
PA   (ARCH-) Arch Development Corp.
PI   Liao S, Chang C;
DR   WPI; 89-309501/42.
DR   N-PSDB; N91578.
PT   New DNA encoding new androgen receptor and TR2 polypeptide(s) - able
PT   to bind DNA, and derived antibodies, useful for receptor assay and
PT   purification.
PS   Claim 8; fig. 3; 60pp; English.
CC   The protein is used to raise antibodies for receptor assays and for
CC   affinity purification.
CC   The 98 kD product starts at the first Met codon; the 79 kD product
CC   starts from the second.
SQ   Sequence   902 AA;

  Query Match             17.8%;  Score 556;  DB 1;  Length 902;
  Best Local Similarity   33.1%;  Pred. No. 2.57e-42;
  Matches   87;  Conservative   73;  Mismatches   92;  Indels   11;  Gaps    8;

Db      542 cliqdeasgchyqaltcgsckvffkraaegkqkylcasrndctidkfrrkncpscrlrk 601
            | :|:| ||| |||   :| :|| ||||: :|    |:|:: |:||||| |||:| :|||||
Qy       96 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 155

Db      602 cyeagmtl qarklk --klgnlklqeegenssagsptedpsqkmtvshi--eqyecqpi 655
            ||| ||    |:|: :    :|          : :    || :     ::   |  :  ::   : :
Qy      156 CYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 215

Db      656 flnvlealepgvvcaghdnnqpdsfaallsslnelgerqlvhvvkwakalpgfrnlhvdd 715
             | :||||  | |:   :: :  | : |::: ||  |::::|||:: ||| :||| :| : |
Qy      216 VLTLLEAEPPHVL-ISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 273

Db      716 qmaviqyswmglmvfamgwrsftnvnsrmlyfapdlvfn-eyrmhksrmysqcvrmrhls 774
            |: :::  || ::: :: |||: :   :::: ||||||:: :   :      :         :   :
Qy      274 QVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDMLLATT 331

Db      775 qefgwlqitpqeflcmkalllfs 797
              |  | :   |:||:||::|::
Qy      332 SRFRELKLQHKEYLCVKAMILLN 354


RESULT   12
ID   P93110 standard; protein; 902 AA.
AC   P93110;
DT   19-MAR-1990  (first entry)
DE   Rat androgen receptor.
KW   Rat androgen receptor, monoclonal antibody, polyclonal antibody, cancer
OS   Rattus rattus.
PN   WO8909791-A.
PD   19-OCT-1989.
PF   13-APR-1989; U01548.
PR   14-APR-1988; US-182646.
PA   (UYNC-) University of North Carolina
PI   French FS, Wilson EM, Joseph DR, Lubahn DB;
DR   WPI; 89-324206/44
DR   N-PSDB; N91773.
PT   DNA encoding androgen receptor protein - useful for transforming
PT   eukaryotic hosts for protein expression and subsequent antibody prodn.
PS   Disclosure; Fig. 5; 41pp; English.
CC   Androgen receptor protein (AR) is used to produce mono- or poly-clonal
CC   antibodies.  These are used for the detection and quantification of AR in
CC   the presence of endogenous androgen, as androgen will not interfere with
CC   binding.  They may be used in assays to determine and quantify cellular
CC   distribution of AR in tumour tissue, and are esp. useful for evaluating
CC   prostate cancers to determine responsiveness to androgen withdrawal
CC   therapy.
SQ   Sequence   902 AA.

  Query Match             17.8%;  Score 556;  DB 1;  Length 902;
  Best Local Similarity   33.1%;  Pred. No. 2.57e-42;
  Matches   87;  Conservative   73;  Mismatches   92;  Indels   11;  Gaps    8;

Db      542 cliqdeasgchyqaltcgsckvffkraaegkqkylcasrndctidkfrrkncpscrlrk 601
            | :|:| ||| |||   :| :|| ||||: :|    |:|:: |:||||| |||:| :|||||
Qy       96 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCIIDKNRRKSCQACRLRK 155

Db      602 cyeagmtl qarklk -klgnlklqeegenssagsptedpsqkmtvshi--eqyecqpi 655
            ||| ||    |:|: :    :|          : :    || :     ::   |  :  ::   : :
Qy      156 CYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 215

Db      656 flnvlealepgvvcaghdnnqpdsfaallsslnelgerqlvhvvkwakalpgfrnlhvdd 715
             | :||||  | |:   :: :  | : |::: ||  |::::|||:: ||| :||| :| : |
Qy      216 VLTLLEAEPPHVL-ISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 273

Db      716 qmaviqyswmglmvfamgwrsftnvnsrmlyfapdlvfn-eyrmhksrmysqcvrmrhls 774
            |: ...  || ... .. |||. .   ...: ||||||:: :   :      :         :   :
Qy      274 QVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDMLLATT 331

Db      775 qefgwlqitpqeflcmkalllfs 797
              |  | :   |:||:||::|::
Qy      332 SRFRELKLQHKEYLCVKAMILLN 354


RESULT   13
ID   R12223 standard; Protein; 918 AA.
AC   R12223;
DT   20-AUG-1991  (first entry)
DE   Human androgen receptor.
KW   hAR; DNA-binding protein; steroid hormone.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   Domain          556..626
FT   /label= DNA-binding domain
FT   /note- "cysteine-rich"
PN   WO9107423-A.
PD   30-MAY-1991.
PF   19-OCT-1990; U06015.
PR   17-NOV-1989; US-438775.
PA   (ARCH-) ARCH DEV CORP.
PI   Liao S, Chang C;
DR   WPI; 91-178048/24.
DR   N-PSDB; Q12001.
PT   Androgen receptor and TR2 DNA binding proteins - DNA sequences
PT   and antibodies for detection and quantification methods
PS   Claim 25; Fig 3; 79pp; English.
CC   This sequence was deduced from a cDNA clone isolated by screening
CC   commercially available human testis and prostate lambda gt11 cDNA
CC   libraries  The sequence is very similar to that of rat AR and in
CC   the DNA-binding domain it is identical to that of rAR DNA-binding
CC   domain. Homology comparisons with other known steroid receptors
CC   indicate that hAR is more closely related to glucocorticoid,
CC   mineralo-corticoid and progesterone receptors than to vitamin A or to
CC   receptors for oestrogen, vitamin D and thyroid hormones.
SQ   Sequence   918 AA;

  Query Match             17.8%;  Score 556;  DB 3;  Length 918;
  Best Local Similarity   32.7%;  Pred. No. 2.57e-42;
  Matches   86;  Conservative   76;  Mismatches   90;  Indels   11;  Gaps    9;

Db      558 cliqdeasgchyqaltcgsckvffkraaegkqkylcasrndctidkfrrknapscrlrk 617
            | :|:| ||| |||   :| :|| ||||: :|    |:|:: |:||||| |||:| :|||||
Qy       96 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCIIDKNRRKSCQACRLRK 155

Db      618 cyeagmtl-qarklk---klgn-lklqeeg-easstsptcettqkltvshieqyecqpi 671
            ||| ||    |:|: :    :|       :   :|       :: :   :     : ::       :::     : :
Qy      156 CYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 215

Db      672 flnvlealepgvvcaghdnnqpdsfaallsslnelgerqlvhvvkwakalpgfrnlhvdd 731
             | .|||  | |.   .. . .  | . |::: ||  |::::|||:: ||| :||| :| : |
Qy      216 VLTLLEAEPPHVL-ISRPSA-PFIEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 273

Db      732 qmavieyswmglmvfamgwrsftnvnsrmlyfapdlvfn-eyrmhksrmysqcvrmrhls 790
            |. ..|  || ... .. |||. .   .... ||||||:: :   :      :         :   :
Qy      274 QVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDMLLATT 331
```

```
Db     791 qefgwlqitpqeflcmkalllfs 813
                |  | :    |:|::||:::::
Qy     332 SRFRELKLQHKEYLCVKAMILLN 354


RESULT   14
ID   P90996 standard; protein; 919 AA.
AC   P90996;
DT   28-FEB-1990  (first entry)
DE   Human androgen receptor DNA clone
KW   Androgen receptor; TR2 polypeptide;
OS   Homo sapiens.
FH   Key                 Location/Qualifiers
FT   Region              1..919
FT   /*tag=  a
FT   /product=98 kD polypeptide
FT   Region              185..919
FT   /*tag=  b
FT   /product=79 kD polypeptide
PN   WO8909223-A.
PD   05-OCT-1989.
PF   24-MAR-1989; U01238.
PR   30-MAR-1988; US-176107.
PA   (ARCH-) Arch Development Corp.
PI   Liao S, Chang C;
DR   WPI; 89-309501/42.
DR   N-PSDB; N91577.
PT   New DNA encoding new androgen receptor and TR2 polypeptide(s) - able
PT   to bind DNA  and derived antibodies  useful for receptor assay and
PT   purification.
PS   Claim 8; fig  3; 60pp; English
CC   The protein is used to raise antibodies for receptor assays and for
CC   affinity purification.
CC   The 98 kD product starts at the first Met codon; the 79 kD product
CC   starts from the second
SQ   Sequence   919 AA;

  Query Match             17.7%;  Score 553;  DB 1;  Length 919;
  Best Local Similarity   32.3%;  Pred. No. 5.10e-42;
  Matches    85;  Conservative   77;  Mismatches  90;  Indels  11;  Gaps    9;

Db     559 cliqgdeasgchygaltcgsckvffkraaegkqkylcasrndctidkfrrknepscrlrk 618
           | :|:| ||| |||   :| .|| ||||. .|    |.|.. |.||||| |||.| :|||||
Qy      96 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRLRK 155

Db     619 cyeagmtl-qarklk---klgn-lklqeeg-easttsptettqkltvshiegyecqpi 672
           ||| ||    |:|.  .    .|      .  .|     :: :  :    : ::     ::: :  :
Qy     156 CYEVGMVKGGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 215

Db     673 flnvleaiepgvvcaghdnnqpdsfaallsslnelgerqlvhvvkwakalpgfrnlhvdd 732
             :||| | |:  :: :  | : |::: || |:...|||:: ||| :||| :| : |
Qy     216 VLTLLEAEPPHVL-ISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 273

Db     733 qmaviqyswmglmvfamgwrsftnvnsrmlyfapdlvfn-eyrmhksrmysqcvrmrhls 791
           |: :::  || ::: :: |||: :  :::: |||||::: :  :    :       :   :
Qy     274 QVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDMLLATT 331

Db     792 qefgwlqitpqeflcmkalllfs 814
                |  | :    |:||:||::|::
Qy     332 SRFRELKLQHKEYLCVKAMILLN 354


RESULT   15
ID   W14783 standard; Protein; 919 AA.
AC   W14783;
DT   22-JUN-1997  (first entry)
DE   Androgen receptor.
KW   Androgen receptor; acidic fibroblast growth factor, aFGF,
KW   antisense; benign prostatic hyperplasia; prostate cancer; therapy.
OS   Homo sapiens.
PN   WO9711170-A1.
PD   27-MAR-1997.
PF   20-SEP-1996; U15081.
PR   20-SEP-1995; US-004018.
PA   (WORC-) WORCESTER FOUND BIOMEDICAL RES.
PI   Zamecnik PA;
DR   WPI; 97-202879/18.
DR   N-PSDB; T63407.
PT   Oligonucleotide(s) antisense to human androgen receptor and acidic
PT   FGF genes - used to inhibit gene expression, for the treatment of
PT   benign prostatic hyperplasia
PS   Disclosure; Page 22-28; 51pp; English.
CC   Human androgen receptor (W14783) binds testosterone and, acting
CC   at the transcriptional level, regulates the growth of normal
CC   prostatic cells  Antisense oligonucleotides (see also T63400,
CC   T63404-05) based on an androgen receptor cDNA clone (see also
CC   T63407) can be used to prevent androgen receptor gene expression,
CC   thereby inhibiting the growth or survival of prostatic cells for
CC   the treatment of benign prostatic hyperplasia and prostate cancer.
SQ   Sequence   919 AA;

  Query Match             17.7%;  Score 553;  DB 21;  Length 919;
  Best Local Similarity   32.3%;  Pred. No. 5.10e-42;
  Matches    85;  Conservative   77;  Mismatches  90;  Indels  11;  Gaps    9;

Db     559 cliqgdeasgchygaltcgsckvffkraaegkqkylcasrndctidkfrrknepscrlrk 618
           | :|:| ||| ||   :| :|| ||||: :|    |:|:: |:||||| |||:| :|||||
Qy      96 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRLRK 155

Db     619 cyeagmtl-qarklk---klgn-lklqeeg-easttsptettqkltvshiegyecqpi 672
           ||| ||    |.|: :    :|      :  :|     :: :  :    : ::     ::: :  :
Qy     156 CYEVGMVKGGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 215

Db     673 flnvleaiepgvvcaghdnnqpdsfaallsslnelgerqlvhvvkwakalpgfrnlhvdd 732
             | :|||  | |            |   :     ||  |....|||.: ||| :||| :| : |
Qy     216 VLTLLEAEPPHVL-ISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 273

Db     733 qmaviqyswmglmvfamgwrsftnvnsrmlyfapdlvfn-eyrmhksrmysqcvrmrhls 791
           |: :::  || ::: :: |||: :  :::: |||||||:: :  :     :       :   :
Qy     274 QVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDMLLATT 331

Db     792 qefgwlqitpqeflcmkalllfs 814
                |  | .    |.||.||..|::
Qy     332 SRFRELKLQHKEYLCVKAMILLN 354


Search completed: Thu Dec 18 09.04.27 1997
Job time : 80 secs.
```

```
      *******************************************************************
       __   __   _____    _____   _____   _____   _   _
      |  \ /  | |  __ \  / ____| |  __ \ / ____| | | | |
      |   V   | | |__) || (___   | |__) | |      | |_| |
      | |\_/| | |  ___/  \___ \  |  _  /| |      |  _  |
      | |   | | | |      ____) | | | \ \| |____  | | | |
      |_|   |_| |_|     |_____/  |_|  \_\\_____| |_| |_|  (TM)

      *******************************************************************

           Release 2.1I John F. Collins, Biocomputing Research Unit
           Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
                   Distribution rights by IntelliGenetics, Inc.

MPsrch_pp   protein - protein database search, using Smith-Waterman algorithm

  n on:          Thu Dec 18 03:08:07 1997,  MasPar time 4.85 Seconds
                                            417.146 Million cell updates/sec
 ular output not generated.

Title:           >US-08-826-361A-21
Description:     (1-418) from US08826361A.pep
Perfect Score:   3116
Sequence:        1 MNYSIPSNVTNLEGGPGRQT          SMPLANLLMLLSHVRHARSA 418

Scoring table:   PAM 150
                 Gap 11

Searched:        54101 seqs, 4836506 residues

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database:        a-issued
                   1:backl 2:51 3:52 4:53 5:54 6:55 7:56 8:PCT90 9:PCT91
                   10:PCT92 11:PCT93 12:PCT94 13:PCT95 14:PCT96

Statistics       Mean 32.531   Variance 144.315,   scale 0.225

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                   SUMMARIES
                  %
 sult           Query
  No.   Score   Match Length DB  ID            Description               Pred. No.
------------------------------------------------------------------------------------
    1    1428    45.8    410  1  5223606-5     Patent No. 5223606.       4.91e-123
    2     660    21.2    363  1  5223606-6     Patent No. 5223606.       5.81e-50
    3     609    19.5    356  1  5223606-7     Patent No. 5223606.       3.45e-45
    4     577    18.5    795  3  US-07-716-    Sequence 5, Applicatio    3.33e-42
    5     505    16.2    446  5  US-07-952-    Sequence 4, Applicatio    1.61e-35
    6     505    16.2    533  5  US-07-952-    Sequence 1, Applicatio    1.61e-35
    7     494    15.9    467  9  PCT-US91-0    Sequence 4, Applicatio    1.68e-34
    8     493    15.8    462  9  PCT-US91-0    Sequence 2, Applicatio    2.08e-34
    9     461    14.8    463  9  PCT-US91-0    Sequence 6, Applicatio    1.85e-31
   10     420    13.5    440  6  US-08-333-    Sequence 8, Applicatio    1.07e-27
   11     411    13.2    454  1  5260432-2     Patent No. 5260432.       7.07e-27
   12     407    13.1    368  1  5223606-3     Patent No. 5223606.       1.64e-26
   13     407    13.1    448  1  5223606-2     Patent No. 5223606.       1.64e-26
   14     393    12.6    462 10  PCT-US92-0    Sequence 4, Applicatio    3.09e-25
   15     393    12.6    797 10  PCT-US92-0    Sequence 2, Applicatio    3.09e-25
   16     378    12.1    462  1  5171671-2     Patent No. 5171671.       7.12e-24
   17     361    11.6    469 13  PCT-US95-1    Sequence 2, Applicatio    2.47e-22
   18     357    11.5     89  6  US-08-561-    Sequence 3, Applicatio    5.67e-22
   19     357    11.5     89  5  US-08-087-    Sequence 3, Applicatio    5.67e-22
   20     353    11.3    494 13  PCT-US95-1    Sequence 1, Applicatio    1.30e-21
   21     349    11.2     66  1  5217867-1     Patent No. 5217867.       2.99e-21
   22     329    10.6    451 13  PCT-US95-1    Sequence 2, Applicatio    1.87e-19
```

```
   23     315    10.1     70  3  US-07-737-    Sequence 3, Applicatio    3.36e-18
   24     307     9.9    443  7  US-08-342-    Sequence 4, Applicatio    1.73e-17
   25     307     9.9    446 13  PCT-US95-1    Sequence 3, Applicatio    1.73e-17
   26     311    10.0    461 13  PCT-US95-1    Sequence 2, Applicati     [illegible]
   27     311    10.0    461 13  PCT-US95-1    Sequence 1, Applicati     [illegible]
   28     310     9.9    461  7  US-08-330-    Sequence 1, Applicatio    [illegible]
   29     310     9.9    461  7  US-08-330-    Sequence 2, Applicatio    9.37e-18
   30     309     9.9    548  6  US-08-333-    Sequence 2, Applicatio    1.15e-17
   31     304     9.8    513  7  US-08-464-    Sequence 2, Applicatio    3.21e-17
   32     304     9.8    558  6  US-08-333-    Sequence 12, Applicati    3.21e-17
   33     303     9.7    410  1  5438126-2     Patent No. 5438126.       3.94e-17
   34     299     9.6    460  7  US-08-342-    Sequence 2, Applicatio    8.93e-17
   35     300     9.6    601  6  US-08-333-    Sequence 14, Applicati    7.28e-17
   36  [illegible]       355  1  5223606-4     Patent No. 5223606.       [illegible]
   37     282     9.1    433 12  PCT-US94-0    Sequence 2, Applicatio    2.86e-15
   38     265     8.5    439  6  US-08-333-    Sequence 10, Applicati    8.98e-14
   39     238     7.6     99  5  US-08-087-    Sequence 2, Applicatio    2.02e-11
   40     238     7.6     99  6  US-08-561-    Sequence 2, Applicatio    2.02e-11
   41     191     6.1     60 12  PCT-US94-1    Sequence 1, Applicatio    2.07e-07
   42     169     5.4    367  3  US-07-737-    Sequence 4, Applicatio    1.39e-05
   43     141     4.5    272 13  PCT-US95-0    Sequence 22, Applicati    2.52e-03
   44     141     4.5    272  6  US-08-485-    Sequence 23, Applicati    2.52e-03
   45      83     2.3    427  1  5310667-7     Patent No. 5310667.       1.94e-01

                                   ALIGNMENTS

RESULT     1
ID   5223606-5     STANDARD;      PRT;   444 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Patent No. 5223606.
XX
CC   Patent No. 5223606
CC        APPLICANT:  BLAUDIN DE THE, HUGHES,MARCHIO, AGNES,TIOLLAIS,
CC   PIERRE;DEJEAN, ANNE
CC        TITLE OF INVENTION: STEROID/THYROID HORMONE RECEPTOR-RELATED
CC   PROTEIN INAPPROPRIATELY EXPRESSED IN HUMAN HEPATOCELLULAR CARCINOMA
CC        NUMBER OF SEQUENCES: 11
CC        CURRENT APPLICATION DATA:
CC          APPLICATION NUMBER:  US/07/134,130
CC          FILING DATE: 17-DEC-1987
CC        PRIOR APPLICATION DATA:
CC   SEQ ID NO:5:
CC          LENGTH: 410
SQ   SEQUENCE   444 AA;  50020 MW;  1110634 CN;

  Query Match             45.8%;  Score 1428,  DB 1,  Length 410;
  Best Local Similarity   57.7%;  Pred. No. 4.91e-123;
  Matches   191,  Conservative   78,  Mismatches  50,  Indels  12;  Gaps    9;

Db          1 CAVCNDYASGYHYGVWSCEGCKAFFKRSIQGHNDYMCPATNQCTIDKNRRKSCQACRLRK 60
              ||||.|||||||||||||||||||||||||||||||.|||||||||||||||||||||||
Qy         96 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 155

Db         61 CYEVGMMKGGIRKDRRGGRMLKHKRQRDDGEGRGEVGSAGDMRAANLWPSPLMIKRSKKN 120
              ||||||:| | |  |.| | |:..  |||    | :   :. . .|..     .|  : |
Qy        156 CYEVGMVKCGSRRERCGYRLVR--RQRSADE-QLHCAGKA-KRSGG--HAP-RV-RELL- 206

Db        121 SLALSLTADQMVSALLDAEPPILYSEYDPTRPFSEASMMGLLTNLADRELVHMINWAKRV 180
               |  .:|.::|:| :||:|||          |: ||:|||||  ||:||:||||||:|||::.
Qy        207  LD-ALSPEQLVLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIP 264

Db        181 GFVDLTLHDQVHLLECAWLEILMIGLVWRSMEHPVKLLFAPNLLLDRNQGKCVEGMVEIF 240
              |||.|.| |||:|||  |:|:|||:|||:|||:|| ||:|||:|:|||::|||||||::|||
Qy        265 GFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIF 324

Db        241 DMLLATSSRFRMMNLQGEEFVCLKSIILLNSGVYTFLSSTLKSLEEKDHIHRVLDKITDT 300
              ||||||.||||  ..||   |..|.|..||||||..| ...:||   :  :      : ::|: |||:
Qy        325 DMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSRKLAHLLNAVTDA 383
```

```
Db      301 LIHLMAKAGLTLQQQHQRLAQLLLILSHIRH 331
            |: ::||:|:: |||  ||| ||::|||:||
Qy      384 LVWVIAKSGISSQQQSMRLANLLMLLSHVRH 414


RESULT     2
ID   5223606-6      STANDARD;      PRT;   393 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Patent No. 5223606.
XX
CC   Patent No. 5223606
CC        APPLICANT.  BLAUDIN DE THE, HUGHES,MARCHIO, AGNES;TIOLLAIS,
CC   PIERRE;DEJEAN, ANNE
CC        TITLE OF INVENTION. STEROID/THYROID HORMONE RECEPTOR-RELATED
     PROTEIN INAPPROPRIATELY EXPRESSED IN HUMAN HEPATOCELLULAR CARCINOMA
          NUMBER OF SEQUENCES: 11
          CURRENT APPLICATION DATA:
CC          APPLICATION NUMBER:  US/07/134,130
CC          FILING DATE: 17-DEC-1987
CC        PRIOR APPLICATION DATA:
CC   SEQ ID NO:6:
CC          LENGTH: 363
SQ   SEQUENCE   393 AA; 44638 MW; 907537 CN;

  Query Match             21.2%;  Score 660;  DB 1;  Length 363;
  Best Local Similarity   32.2%;  Pred. No. 5.81e-50;
  Matches  104;  Conservative   88;  Mismatches 120;  Indels  11;  Gaps  11;

Db        1 CLICGDEASGCHYGVLTCGSCKVFFKRAMEGQHNYLCAGRNDCIVDKIRRKNCPACRLRK 60
            | :|:| ||| |:|| :| :|| ||||:::|:::|:|:: |:| :|| |||:| ||||||
Qy       96 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 155

Db       61 CCQAGMVL-GGRKFK-KFNKVRVMRALDAVALPAPVGIPNESQRITFSPSQEIQLIP-PL 117
            | : |||  |:|: :  :  || |: |     | : : ::             | |  |
Qy      156 CYEVGMVKGGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 215

Db      118 INLLMSIEPDVIYAGHDNTKPDTSSSLLTSLNQLGERQLLSVVKWSKSLPGFRNLHIDDQ 177
            :  |:  ||  :  :: :: | | :|:  || | : |     | | ||| :   ||
Qy      216 VLTLLEAEPPHVLISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFDQ 274

Db      178 ITLIQYSWMSLMVFGLGWRSYKHVSGQMLYFAPDLILNEQRMK-ESSFYSLCLTMWQIPQ 236
            : |::  || ::: || |||  | :| :: |||||:|: :  |   ::  :   :
        275 VRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDMLLATTS 332

        237 EFVKLQVSQEEFLCMKVLLLLNT-I-PLEGLRSQSQFEEMRSSYIRELIKAIG-LRQKGV 293
             |  | : : |:||:| ::|||: : ||      :::   : :  :  : |:  : |:
Qy      333 RFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKSG 392

Db      294 VSSSQRFYQLTKLLDNLHDLVKQ 316
            :|| |:  :|::||  |   |::
Qy      393 ISSQQQSMRLANLL-MLLSHVRH 414


RESULT     3
ID   5223606-7      STANDARD;      PRT;   386 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Patent No. 5223606.
XX
CC   Patent No. 5223606
CC        APPLICANT:  BLAUDIN DE THE, HUGHES;MARCHIO, AGNES;TIOLLAIS,
CC   PIERRE;DEJEAN, ANNE
CC        TITLE OF INVENTION: STEROID/THYROID HORMONE RECEPTOR-RELATED
CC   PROTEIN INAPPROPRIATELY EXPRESSED IN HUMAN HEPATOCELLULAR CARCINOMA
CC        NUMBER OF SEQUENCES: 11
CC        CURRENT APPLICATION DATA:
CC          APPLICATION NUMBER:  US/07/134,130
CC          FILING DATE: 17-DEC-1987
CC        PRIOR APPLICATION DATA:
CC   SEQ ID NO:7:
CC          LENGTH: 356
SQ   SEQUENCE   386 AA; 44008 MW; 864027 CN;

  Query Match             19.5%;  Score 609;  DB 1;  Length 356;
  Best Local Similarity   36.0%,  Pred. No. 3.45e-45;
  Matches   95,  Conservative   70,  Mismatches  84;  Indels  15;  Gaps  14;

Db        1 CLVCSDEASGCHYGVLTCGSCKVFFKRAVEGQHNYLCAGRNDCIIDKIRRKNCPACRYRK 60
            | |||| ||| |||| :| :|| ||||:::|:::|:|:: |:| ||| |||:| ||| ||
Qy       96 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 155

Db       61 CLQAGMNL-EARKIK-K-KIKGIQQAIPG-V-SQETSENPGNKI-IVPAILPQ-LTPT-L 112
            | : ||    :|: :   ::    |:::    :        :  :|  :  |   | : : :
Qy      156 CYEVGMVKGGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 215

Db      113 V-SLLEVIEPEVLYAGTDSSVPDSTWRIMTTLNMLGGRQVIAAVKWAKAIPGFRNLHLDD 171
            | :|||    | || :   ::  :::   :| :| :|: ::::  : ||| |||| :| | |
Qy      216 VLTLLEAEPPHVLISRPSAPFTEAS-MMMSLTKLADKELVHMISWAKKIPGFVELSLFD 273

Db      172 QNTLLQYSWM-FLMAALGWRSYRQSSANLLCFAPDLIIN-EQRMILPCMYDQCKHMLYVS 229
            |  ||:  ||  || :| |||  ...  .|: |||||:.. .. .  .  .  .   :|  :
Qy      274 QVRLLESCWMEVLMMGLMWRSIDHPG-KLI-FAPDLVLDRDEGKCVEGILEIFDMLLATT 331

Db      230 SELHRLQVSYEEYLCMKYLLLLSS 253
            | :: | :    ||||:| ::||:|
Qy      332 SRFRELKLQHKEYLCVKAMILLNS 355


RESULT     4
ID   US-07-716-827C-5      STANDARD;      PRT;   795 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 5, Application US/07716827C.
XX
CC    Sequence 5, Application US/07716827C
CC    Patent No. 5215916
CC     GENERAL INFORMATION:
CC       APPLICANT:  Simons Jr., Stoney S.
CC       APPLICANT:  Yamamoto, K. R.
CC       APPLICANT:  Chakraborti, P. K.
CC       APPLICANT:  Garabedian, M. J.
CC       TITLE OF INVENTION.  SUPER GLUCOCORTICOID RECEPTORS
CC       NUMBER OF SEQUENCES:  5
CC       CORRESPONDENCE ADDRESS.
CC         ADDRESSEE:  Cushman, Darby & Cushman
CC         STREET.  Eleventh floor, 1615 L Street, N.W.
CC         CITY:  Washington
CC         STATE:  D.C.
CC         COUNTRY:  U.S.A.
CC         ZIP:  20036-5601
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/716,827C
CC         FILING DATE:  19910619
CC         CLASSIFICATION:  435
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Scott, Watson T.
CC         REGISTRATION NUMBER:  26,581
CC         REFERENCE/DOCKET NUMBER:  WTS/5683/84453
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (202)8612000
CC         TELEFAX:  (202)822-8944
CC         TELEX:  6714627 CUSH
CC     INFORMATION FOR SEQ ID NO:  5:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  795 amino acids
CC         TYPE:  AMINO ACID
CC         TOPOLOGY:  linear
SQ   SEQUENCE   795 AA; 87498 MW; 3326890 CN;

  Query Match             18.5%;  Score 577;  DB 3;  Length 795;
  Best Local Similarity  32.9%;  Pred. No. 3.33e-42;
  Matches   92;  Conservative   83;  Mismatches  83;  Indels  16;  Gaps  15;

Db    424 SPPSSSAATIPFFKLTLVTSCEASSTHYGVLTCESCKVFFKRAVEGQHNYLCAGRNDCI 483
          ::| .....      . .| |||| ||| |||| .| .|| ||||...|::::|::: |:|
Qy     81 ASPVTGPGSK-RDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCT 139

Db    484 IDKIRRKNCPACRYRKCLQAGMNI-EAPKTK-K-KIKCIQQATAG-V-SQDTSENPN-KT 537
          ||| ||||:| ||| ||| : ||     :|: :    ..   |.
      140 IDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQRSALEQLHCAGKAKRSGGHA 199

      538  IVPAAIPQ-LTPT-IV-SLLFVIEPEVLYAGYDSSVPDSAWRIMTTLNMLGGRQVIAAV 593
           |  : : |:|  || :|||   | ||                   :|  |  |.
Qy    200 PRVRELLLDALSPEQLVLTLLEAEPPHVLISRPSAPFTEAS--MMMSLTKLADKELVHMI 257

Db    594 KWAKAIPGLRNLHLDDQMTLLQYSWMFLMAFALGWRSYRQSSGNLLCFAPDLIIN-EQRM 652
           ||| | :: :| : ||. ||.  ||  ..  .| |:|  .. |.|. |||||... .. .
Qy    258 SWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHP-GKLI-FAPDLVLDRDEGK 315

Db    653 SLPCMYDQCKHMLFVSSELQRLQVSYEEYLCMKTLLLLSS 692
           :  .  .  .| .| . .. | .  ||||.|.  ||:|
Qy    316 CVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNS 355


RESULT     5
ID   US-07-952-800-4     STANDARD;      PRT;   446 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 4, Application US/07952800
XX
CC    Sequence 4, Application US/07952800
CC    Patent No. 5403925
CC     GENERAL INFORMATION:
CC       APPLICANT:  OZATO, KEIKO
CC       TITLE OF INVENTION:  A NEW MEMBER OF THE NUCLEAR HORMONE
CC       TITLE OF INVENTION:  RECEPTOR SUPERFAMILY AND A CDNA CLONE THEREOF
CC       NUMBER OF SEQUENCES:  4
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  TOWNSEND AND TOWNSEND
CC         STREET:  1 MARKET PLAZA, STEUART TOWER, SUITE 2000
CC         CITY:  SAN FRANCISCO
CC         STATE:  CA
CC         COUNTRY:  USA
CC         ZIP:  94105
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/952,800
CC         FILING DATE:  19920928
CC         CLASSIFICATION:  435
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Weber, Kenneth A
CC         REGISTRATION NUMBER:  31,677
CC         REFERENCE/DOCKET NUMBER:  15280-21-1
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  415-543-9600
CC         TELEFAX:  415-543-5043
```

```
CC     INFORMATION FOR SEQ ID NO:  4:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  446 amino acids
CC         TYPE:  AMINO ACID
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   446 AA; 48023 MW; 964846 CN;

  Query Match             16.2%;  Score 505;  DB 5;  Length 446;
  Best Local Similarity  36.2%;  Pred. No. 1.61e-35;
  Matches  102;  Conservative   63;  Mismatches  98;  Indels  19;  Gaps  13;

Db     99 VRGIHCPPPPGGPGAGK--RLCAICGDRSSGKHYGVYSCEGCKGFFKRTIRKDLIYSCRD 156
          | | .|..|  |||. .   ..||.|.| .|| |||| |||||.||||.|:     | |
Qy     75 VSGNRCASPVTGPGSKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPA 134

Db    157 NKDCIVDKRQRNRCQYCRYQKCLAIGM-K----RE--AVQ-EERQRGKDKDGDGDG-AGG 207
           ..:||| :|. || || .||   || |    ||  .  .  |||: | :    . |
Qy    135 INQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQRSALEQLHCAGKAKR 194

Db    208 APEEMPVDR-ILEAELAVEQKS-DQ-GVEGPGATGGGGSSP-NDP--VTNICQAADKQLF 261
          :   |  | .|   |. ||          ||    . |:|  ..  : ::   ||||:|
Qy    195 SGGHAPRVRELLLDALSPEQLVLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELV 254

Db    262 TLVEWAKRIPHFSSLPLDDQVILLRAGWNELLIASFSHRSIDVRDGILLATGLHVHRNSA 321
            .. |||.||| |  |:| ||| ||  .  | |.|: ::  ||||     :::|  | : .: :
Qy    255 HMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEG 314

Db    322 HSA-GVGAIFDRVLTELVSKMRDMRMDKTELGCLRAIILFNP 362
              |.  ||| .|.   |. |.....  |  |..|:|:|:
Qy    315 KCVEGILEIFDMLLAI-ISRFRELKLQHKEYLCVKAMILLNS 355


RESULT     6
ID   US-07-952-800-2     STANDARD;      PRT;   533 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 2, Application US/07952800.
XX
CC    Sequence 2, Application US/07952800
CC    Patent No. 5403925
CC     GENERAL INFORMATION:
CC       APPLICANT:  OZATO, KEIKO
CC       TITLE OF INVENTION:  A NEW MEMBER OF THE NUCLEAR HORMONE
CC       TITLE OF INVENTION:  RECEPTOR SUPERFAMILY AND A CDNA CLONE THEREOF
CC       NUMBER OF SEQUENCES:  4
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  TOWNSEND AND TOWNSEND
CC         STREET:  1 MARKET PLAZA, STEUART TOWER, SUITE 2000
CC         CITY:  SAN FRANCISCO
CC         STATE:  CA
CC         COUNTRY:  USA
CC         ZIP:  94105
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/952,800
CC         FILING DATE:  19920928
CC         CLASSIFICATION:  435
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Weber, Kenneth A
CC         REGISTRATION NUMBER:  31,677
CC         REFERENCE/DOCKET NUMBER:  15280-21-1
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  415-543-9600
CC         TELEFAX:  415-543-5043
CC     INFORMATION FOR SEQ ID NO:  2:
```

```
CC         SEQUENCE CHARACTERISTICS:
CC           LENGTH:  533 amino acids
CC           TYPE:  AMINO ACID
CC           TOPOLOGY:  linear
CC         MOLECULE TYPE:  protein
SQ   SEQUENCE   533 AA, 56921 MW, 1406341 CN,

  Query Match              16.2%;  Score 506;  DB 5;  Length 533;
  Best Local Similarity   36.2%,  Pred. No. 1.30e-35;
  Matches   102;  Conservative   63;  Mismatches  98;  Indels  19;  Gaps  13;

Db      186 VRGLHCPPPPGGPGAGK--ELCAICGDRSSGKHYGVYSCEGCKGFFKRTIRKDLTYSCRD 243
             | |  :|::|  |||: :   ::||:|:| :|| |||| ||||||:||||.|.     | |
Qy       75 VSGNRCASPVTGPGSKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPA 134

Db      244 NKDCTVDKRQRNRCQYCRYQKCLATGM-K----RE--AVQ-EERQRGKDKDGDGEG-AGG 294
              ::||:|| ::: || || :||    |||     ||  . .    |||. | :     | |
Qy      135 TNQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKR 194

Db      295 APEEMPVDR-ILEAELAVEQKS-DQ-GVEGPGGTGGSGSSP-NDP--VTNICQAADKQLF 348
             :     |   | :|    |: ||                | |      : |:| :: : ::  |||:|
Qy      195 SGGHAPRVRELLLDALSPEQLVLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELV 254

Db      349 ILVEWAKRIPHFSSLPLDDQVILLRAGWNELLIASFSHRSIDVRDGILLATGLHVHRNSA 408
              :: |||:|| |   |:| ||| || : | |:|· ··  ||||      ···|  | : |: :
Qy      255 HMISWAKKIPGFVELSLEDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEG 314

Db      409 HSA-GVGAIFDRVLTELVSKMRDMRMDKTELGCLRAIILFNP 449
                 |:  ||| :|:    |: |:::::  |   |::|:||:|:
Qy      315 KCVEGILEIFDMLLAT-TSRFRELKLQHKEYLCVKAMILLNS 355

RESULT     7
ID   PCT-US91-00399-4      STANDARD;      PRT;   467 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 4, Application PC/TUS9100399.
XX
CC    Sequence 4, Application PC/TUS9100399
CC     GENERAL INFORMATION:
CC         APPLICANT:  Mangelsdorf, Dr., David J.
CC         APPLICANT:  Evans Dr., Ronald M
CC         TITLE OF INVENTION:  RETINOID RECEPTOR COMPOSITIONS AND
CC         TITLE OF INVENTION:  METHODS
CC         NUMBER OF SEQUENCES:  7
CC         CORRESPONDENCE ADDRESS:
CC           ADDRESSEE:  FITCH, EVEN, TABIN & FLANNERY
CC           STREET:  135 South LaSalle Street, Suite 900
CC           CITY:  Chicago
CC           STATE:  Illinois
CC           COUNTRY:  USA
CC           ZIP:  60603
CC         COMPUTER READABLE FORM:
CC           MEDIUM TYPE:  Floppy disk
CC           COMPUTER:  IBM PC compatible
CC           OPERATING SYSTEM:  PC-DOS/MS-DOS
CC           SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC         CURRENT APPLICATION DATA:
CC           APPLICATION NUMBER:  PCT/US91/00399
CC           FILING DATE:  19910122
CC           CLASSIFICATION:  435
CC         PRIOR APPLICATION DATA:
CC           APPLICATION NUMBER:  07/478,071
CC           FILING DATE:  09-FEB-1990
CC         ATTORNEY/AGENT INFORMATION:
CC           NAME:  Reiter, Mr., Stephen E.
CC           REGISTRATION NUMBER:  31192
CC           REFERENCE/DOCKET NUMBER:  50852
CC         TELECOMMUNICATION INFORMATION:
CC           TELEPHONE:  (619) 552 1311
```

```
CC           TELEFAX:  (619) 552-0095
CC           TELEX:  20 6566 PATLAW CGO
CC       INFORMATION FOR SEQ ID NO:  4:
CC         SEQUENCE CHARACTERISTICS:
CC           LENGTH:  467 amino acids
CC           TYPE:  AMINO ACID
CC           TOPOLOGY:  linear
CC         MOLECULE TYPE:  protein
SQ   SEQUENCE   467 AA; 51215 MW; 1083042 CN;

  Query Match              15.9%;  Score 494;  DB 9;  Length 467;
  Best Local Similarity   34.1%,  Pred. No. 1.68e-34;
  Matches    95;  Conservative   71;  Mismatches  94;  Indels  19;  Gaps  15;

Db      138 HICAICGDRSSGKHYGVYSCEGCKGFFKRTVRKDLTYTCRDNKDCLIDKRQRNRCQYCRY 197
             |:||:|:| :|| |||| ||||||:||||:::     | |    ::| ||| :|: || '|
Qy       94 HFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRL 153

Db      198 QKCIAMGMKREAVQEER-Q-R-GK-DRNENE-VESTSSAN---EDMP-VEKIIEAELAVE 248
             :||  :|| : : : ||   |  : :|: :| :  :: |:       | |  :|   |: '
Qy      154 RKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPE 213

Db      249 PKTETYVEANMG--L-N-PSSP-NDP--VINICQAADKQLFILVEWAKRIPHFSELPLDD 301
                   | :||:      | : ||:|  ::   : ::    |||:|  :: |||:|| | ||:| |
Qy      214 QLVLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLED 273

Db      302 QVILLRAGWNELLIASFSHRSIAVKDGILLATGLHVHRNSAHSA-GVGAIFDRVLTELVS 360
             || || : | |:|: ::  |||        :::|  | : ': :      |:  ||| :|:     |
Qy      274 QVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLAT-TS 332

Db      361 KMRDMQMDKTELGCLRAIVLFNPDSKGLSNPA-EVEALR 398
             : |:: ::   |   |::|::|:|:       |   :: : :: |
Qy      333 RFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSR 371

RESULT     8
ID   PCT-US91-00399-2      STANDARD;      PRT;   462 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 2, Application PC/TUS9100399.
XX
CC    Sequence 2, Application PC/TUS9100399
CC     GENERAL INFORMATION:
CC         APPLICANT:  Mangelsdorf, Dr., David J.
CC         APPLICANT:  Evans Dr., Ronald M.
CC         TITLE OF INVENTION:  RETINOID RECEPTOR COMPOSITIONS AND
CC         TITLE OF INVENTION:  METHODS
CC         NUMBER OF SEQUENCES:  7
CC         CORRESPONDENCE ADDRESS:
CC           ADDRESSEE:  FITCH, EVEN, TABIN & FLANNERY
CC           STREET:  135 South LaSalle Street, Suite 900
CC           CITY:  Chicago
CC           STATE:  Illinois
CC           COUNTRY:  USA
CC           ZIP:  60603
CC         COMPUTER READABLE FORM:
CC           MEDIUM TYPE:  Floppy disk
CC           COMPUTER:  IBM PC compatible
CC           OPERATING SYSTEM:  PC-DOS/MS-DOS
CC           SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC         CURRENT APPLICATION DATA:
CC           APPLICATION NUMBER:  PCT/US91/00399
CC           FILING DATE:  19910122
CC           CLASSIFICATION:  435
CC         PRIOR APPLICATION DATA:
CC           APPLICATION NUMBER:  07/478,071
CC           FILING DATE:  09-FEB-1990
CC         ATTORNEY/AGENT INFORMATION:
CC           NAME:  Reiter, Mr., Stephen E.
CC           REGISTRATION NUMBER:  31192
CC         REFERENCE/DOCKET NUMBER:  50852
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (619) 552-1311
CC         TELEFAX:  (619) 552-0095
CC         TELEX:  20 6566 PATLAW CGO
CC     INFORMATION FOR SEQ ID NO:  2:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  462 amino acids
CC         TYPE:  AMINO ACID
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   462 AA; 50811 MW; 1063517 CN;

  Query Match             15.8%;  Score 493;  DB 9;  Length 462;
  Best Local Similarity   34.1%;  Pred. No. 2.08e-34;
  Matches    95;  Conservative   71;  Mismatches  94;  Indels  19;  Gaps  15;

Db    133 HICAICGDRSSGKHYGVYSCEGCKGFFKRTVRKDLTYTCRDNKDCLIDKRQRNRCQYCRY 192
          |:||:|:| :|| |||| ||||||.||||            | |       | |||  |   || ||
Qy     94 HFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRL 153

      193 QKCLAMGMKREAVQEER-Q-R-GK-DRNENE-VESTSSAN---EDMPVER-ILEAELAVE 243
          :||  :||  : : : ||    |  : :|: :| :  :: |.         :  :  :  :  :
Qy    154 RKCYEVGMVKEGSRRERCGYRIVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPE 213

Db    244 PKTETYVEANMG--L-N-PSSP-NDP--VTNICQAADKQLFTLVEWAKRIPHFSELPLDD 296
                | :||.     | . ||.|  .:   . . ..    ||:.|    .. |||.|| | || | :
Qy    214 QLVLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 273

Db    297 QVILLRAGWNELLIASFSHRSIAVKDGILLATGLHVHRNSAHSA-GVGAIFDRVLTELVS 355
          |: |: : | :|: ::  |||        :::|  | : |: :      |:  ||| :|:     |
Qy    274 QVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLAT-TS 332

Db    356 KMRDMQMDKTELGCLRAIVLFNPDSKGLSNPA-EVEALR 393
           : |:: ::     |:|::|:|:        |   :: : :: |
Qy    333 RFRELKLQHKEYLCVKAMILLNSSMYPLVIATQDADSSR 371


RESULT     9
ID   PCT-US91-00399-6     STANDARD;     PRT;   463 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 6  Application PC/IUS9100399
XX
CC    Sequence 6, Application PC/IUS9100399
       GENERAL INFORMATION:
         APPLICANT:  Mangelsdorf, Dr., David J.
CC       APPLICANT:  Evans Dr., Ronald M.
CC       TITLE OF INVENTION:  RETINOID RECEPTOR COMPOSITIONS AND
CC       TITLE OF INVENTION:  METHODS
CC       NUMBER OF SEQUENCES:  7
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  FITCH, EVEN, TABIN & FLANNERY
CC         STREET:  135 South LaSalle Street  Suite 900
CC         CITY:  Chicago
CC         STATE:  Illinois
CC         COUNTRY:  USA
CC         ZIP:  60603
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US91/00399
CC         FILING DATE:  19910122
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/478,071
CC         FILING DATE:  09-FEB-1990
```

```
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Reiter, Mr., Stephen E.
CC         REGISTRATION NUMBER:  31192
CC         REFERENCE/DOCKET NUMBER:  50852
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (619) 552-1311
CC         TELEFAX:  (619) 552-0095
CC         TELEX:  20 6566 PATLAW CGO
CC     INFORMATION FOR SEQ ID NO:  6:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  463 amino acids
CC         TYPE:  AMINO ACID
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   463 AA; 50978 MW; 1097932 CN;

  Query Match             14.8%;  Score 461;  DB 9;  Length 463;
  Best Local Similarity   33.4%;  Pred. No. 1.85e-31;
  Matches    97;  Conservative   74;  Mismatches  96;  Indels  23;  Gaps  18;

Db    127 PSTSPGSLVK-HICAICGDRSSGKHYGVYSCEGCKGFFKRTIRKDLIYTCRDNKDCLIDK 185
          | |:|||     |.||.|:| :|| |||| |||||:|||:|:         | |    ::: |||
Qy     83 PVIGPGSKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDK 142

Db    186 RQRNRCQYCRYQKCLVMGMKREAVQEER-Q-R--SRER-AESEAECASSS-HED--MP-V 236
           :|: || || :||  .||  .  .  .  ||     |     |.| |. .  ||. . .       | |
Qy    143 NRRKSCQACRLRKCYEVGMVKEGSRRERCGYRIVRRQRSADEQLHCAGKAKRSGGHAPRV 202

Db    237 -EPILEA---P-LAVEP-KTES-YGDMNVENS--TNDPVT-NICHAADKQLFTLVEWAKR 286
           | :|:|    | |  .     .|.      ..   ..  |. ..   ::    ||':'  :: |||:
Qy    203 RELLLDALSPEQLVLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKK 262

Db    287 IPHFSDLTLEDQVILLRAGWNELLIASFSHRSVSVQDGILLATGLHVHRSSAHS-RGVGS 345
          || | .|.| ||| || . | |.|: ..   ||.       ..:|  | : |  :     |:
Qy    263 IPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILE 322

Db    346 IFDRVLTELVSKMKDMQMDKSELGCLRAIVLFNPDAKGLSNPS-EVEILR 394
          ||| .|.    |. ... :.   |  |..|.:|:|:      |   :: : :: |
Qy    323 IFDMLLAT-TSRFRELKLQHKEYLCVKAMILLNSSMYPLVIATQDADSSR 371


RESULT    10
ID   US-08-333-358-8     STANDARD;     PRT;   440 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 8, Application US/08333358.
XX
CC    Sequence 8, Application US/08333358
CC    Patent No. 5571696
CC     GENERAL INFORMATION:
CC       APPLICANT:  EVANS Ph.D., RONALD M.
CC       APPLICANT:  MANGELSDORF Ph.D., DAVID J.
CC       APPLICANT:  ONG Ms., ESTELITA S.
CC       APPLICANT:  ORO Ph.D., ANTHONY E.
CC       APPLICANT:  BORGMEYER Ph.D., UWE K.
CC       APPLICANT:  GIGUERE Ph.D., VINCENT NMN
CC       APPLICANT:  YAO Mr., TSO-PANG NMN
CC       TITLE OF INVENTION:  NOVEL RECEPTORS
CC       NUMBER OF SEQUENCES:  14
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Pretty, Schroeder, Brueggemann & Clark
CC         STREET:  444 So. Flower St., Suite 2000
CC         CITY:  Los Angeles
CC         STATE:  CA
CC         COUNTRY:  US
CC         ZIP:  90071-2921
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC          SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC        CURRENT APPLICATION DATA:
CC          APPLICATION NUMBER:  US/08/333,358
CC          FILING DATE:
CC          CLASSIFICATION:  435
CC        PRIOR APPLICATION DATA:
CC          APPLICATION NUMBER:  US/07/761,368
CC          FILING DATE:
CC        ATTORNEY/AGENT INFORMATION:
CC          NAME:  Reiter Ph.D., Stephen E.
CC          REGISTRATION NUMBER:  31192
CC          REFERENCE/DOCKET NUMBER:  P31 8936
CC        TELECOMMUNICATION INFORMATION:
CC          TELEPHONE:  (619) 535-9001
CC          TELEFAX:  (619) 535-8949
CC      INFORMATION FOR SEQ ID NO:  8:
CC        SEQUENCE CHARACTERISTICS:
CC          LENGTH:  440 amino acids
CC          TYPE:  amino acid
          TOPOLOGY:  linear
        MOLECULE TYPE:  protein
    SEQUENCE   440 AA; 49883 MW; 957063 CN;

  Query Match             13.5%;  Score 420;  DB 6;  Length 440;
  Best Local Similarity   30.8%;  Pred. No. 1.07e-27;
  Matches    84;  Conservative   68;  Mismatches 104;  Indels   17;  Gaps   15;

Db      80 GPAPKMLGNELCSVCGDKASGFHYNVLSCEGCKGFFRRSVIKGAHYICHSGGHCPMDTYM 139
           ||::|      :|:||:| |||:|| | ||||||:||:||:        |||  :   :| :|
Qy      86 GPGSKR-DAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNR 144

Db     140 RRKCQECRLRKCRQAGMREECVLSEEQ--IRLKKLKRQEEEQAHATSLPPRRSSPPQILP 197
           |: || |||||| : ||     |      |:     || :  |   :|| | :: : | :: :  :
Qy     145 RKSCQACRLRKCYEVGM VKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVR 203

Db     198 QLSPEQLGMIEKLVAAQQQCNRRSFSDRLRVTPWPMAPDPHSREARQQRFAHFTELAIVS 257
           :|   :  |:   | ||  :   :  :   :  :   |  :: :| :: | :  ||      :  :| ||
Qy     204 ELLLDALSP EQLVLTLLEAE -P PHVL IS-RP-SA-PFT-EASM-MMS-LTKLADKE 252

Db     258 VQEIVDFAKQLPGFLQLSREDQIALLKTSAIEVMLLETSRRYNPGSESITFLKDFSYNRE 317
           :  ::    || :|||::||  ||: ||  :   :||:::     :|     :  : |   |:  :|:
Qy     253 LVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRD 312

Db     318 DFAKAGLQV-EFINPIFEFSRAMNELQLNDAEF 349
           : .|      . |..: .. .      || |   |:
Qy     313 E-GKCVEGILEIFDMLLATTSRFRELKLQHKEY 344


    LT   11
    5260432-2      STANDARD;      PRT;   492 AA.
XX
AC    xxxxxx
XX
DT    01-JAN-1900
XX
DE    Patent No. 5260432.
XX
CC    Patent No. 5260432
CC        APPLICANT: TAKAKU, FUMIMARO;ISHIKAWA, TAKASHI;IMAWARI, MICHIO;
CC    EVANS, RONALD M.;UMESONO, KAZUHIKO
CC        TITLE OF INVENTION: HUMAN GAMMA RETINOIC ACID RECEPTOR DNA
CC        NUMBER OF SEQUENCES: 4
CC        CURRENT APPLICATION DATA
CC          APPLICATION NUMBER:  US/07/370,407
CC          FILING DATE: 22-JUN-1989
CC    SEQ ID NO:2:
CC          LENGTH: 454
SQ    SEQUENCE   492 AA; 54527 MW; 1333563 CN;

  Query Match             13.2%;  Score 411;  DB 1;  Length 454;
  Best Local Similarity   29.7%;  Pred. No. 7.97e-27;
  Matches    74;  Conservative   51;  Mismatches 102;  Indels   12;  Gaps   12;
```

```
Db      90 CFVCNDKSSGYHYGVSSCEGCKGFFRRSIQKNMVYTCHRDKNCIINKVTRNRCQYCRLQK 149
           | ||:| :||||||| ||||||:||:|||| :    | |      : | |:|   |: || |||:|
Qy      96 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRLRK 155

Db     150 CFEVGMSKEAVRNDR N K KKKEVKEEGSPDSYELSPQLEELITKVSKAHQETF-PS-L 204
           |:|||| | : | :|      :   :::      :               : :       :|        ::: |   |
Qy     156 CYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 215

Db     205 C-QL-GKYTTNSSADHR-VQL-DLGLWDKFSELATKCIIKIVEFAKRLPGFTGLSIADQI 260
               |      :      :       : :  : ::     :: || |  :: ::  ||::|||  ||: ||:
Qy     216 VLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFDQV 275

Db     261 TLLKAAC-LDILMLRICTRYTPEQDTMTFSDGLTLNRTQ-MHNAGFGPLTDLVFAFAGQL 318
            ||   :|  :::||: :   |              : |:   | |:| : :     |:   : |:::: ::::
Qy     276 RLL-ESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSRF 334

Db     319 LPLEMDDTE 327
             | ::   |
Qy     335 RELKLQHKE 343


RESULT   12
ID    5223606-3      STANDARD;      PRT;   398 AA.
XX
AC    xxxxxx
XX
DT    01-JAN-1900
XX
DE    Patent No. 5223606.
XX
CC    Patent No. 5223606
CC        APPLICANT:  BLAUDIN DE THE, HUGHES;MARCHIO, AGNES;TIOLLAIS,
CC    PIERRE;DEJEAN, ANNE
CC        TITLE OF INVENTION: STEROID/THYROID HORMONE RECEPTOR-RELATED
CC    PROTEIN INAPPROPRIATELY EXPRESSED IN HUMAN HEPATOCELLULAR CARCINOMA
CC        NUMBER OF SEQUENCES: 11
CC        CURRENT APPLICATION DATA:
CC          APPLICATION NUMBER:  US/07/134,130
CC          FILING DATE: 17-DEC-1987
CC        PRIOR APPLICATION DATA:
CC    SEQ ID NO:3:
CC          LENGTH: 368
SQ    SEQUENCE   398 AA; 44842 MW; 907111 CN;

  Query Match             13.1%;  Score 407;  DB 1;  Length 368;
  Best Local Similarity   28.9%;  Pred. No. 1.64e-26;
  Matches    72;  Conservative   66;  Mismatches  99;  Indels   12;  Gaps   12;

Db       1 CFVCQDKSSGYHYGVSACEGCKGFFRRSIQKNMIYTCHRDKNCVINKVTRNRCQYCRLQK 60
           | || | .||||||| .|||||:||:|||| :    | |      : | |:|   |: || |||:|
Qy      96 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRLRK 155

Db      61 CFEVGMSKESVRNDR-N-K-KKKETSKQECTESYEMTAELDDLTEKIRKAHQETF-PS-L 115
           |:|||| | : | :|    :   ::: | :|          ::            : ::|     ::: |   |
Qy     156 CYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 215

Db     116 C-QL-GKYTTNSSADHR-VRL-DLGLWDKFSELATKCIIKIVEFAKRLPGFTGLTIADQI 171
               |      .      .       . . .:      :: || |  :: ::  ||::|||  |:: ||:
Qy     216 VLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFDQV 275

Db     172 TLLKAAC-LDILILRICTRYTPEQDTMTFSDGLTLNRTQ-MHNAGFGPLTDLVFTFANQL 229
            ||   :|  :::|:: :   |              : |:   | |:| : :     |:   : |:::: ::::
Qy     276 RLL-ESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSRF 334

Db     230 LPLEMDDTE 238
             | ::   |
Qy     335 RELKLQHKE 343


RESULT   13
ID    5223606-2      STANDARD;      PRT;   485 AA.
XX
AC    xxxxxx
```

```
XX
DT   01-JAN-1900
XX
DE   Patent No. 5223606.
XX
CC   Patent No. 5223606
CC       APPLICANT: BLAUDIN DE THE, HUGHES-MARCHIO  AGNES-TIOLLAIS
CC   PIERRE-DEJEAN, ANNE
CC       TITLE OF INVENTION: STEROID/THYROID HORMONE RECEPTOR-RELATED
CC   PROTEIN INAPPROPRIATELY EXPRESSED IN HUMAN HEPATOCELLULAR CARCINOMA
CC       NUMBER OF SEQUENCES: 11
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER: US/07/134 130
CC         FILING DATE: 17-DEC-1987
CC       PRIOR APPLICATION DATA:
CC   SEQ ID NO:2:
CC         LENGTH: 448
SQ   SEQUENCE   495 AA; 54437 MW; 1345912 CN;

  Query Match             13.1%;  Score 407;  DB 1;  Length 448;
  Best Local Similarity   28.9%;  Pred. No. 1.64e-26;
  Matches   72;  Conservative   66;  Mismatches   44;  Indels   12;  Gaps   12;

Db      81 CFVCGDKSSGYHYGVSACEGCKGFFRRSIQKNMIYTCHRDKNCVINKVTRNRCQYCRLQK 140
           | || |  :||||||: :||||||:||:||||  :   | |    :  | | |:|   |: || ||||:|
Qy      96 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 155

Db     141 CFEVGMSKESVRNDR-N-K-KKKEISKQECTESYEMTAELDDLTEKIRKAHQETF-PS-L 195
           |:|||| |  :  |  :|     :    ::: | :|        ::          : ::|      ::: |  |
Qy     156 CYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 215

Db     196 C-QI-GKYTTNSSADHR-VRL-DLGLWDKFSELATKCIIKIVEFAKRLPGFTGLTIADQI 251
                 |        :      :        :  :  ::       :: || | :: ::    ||::|||  |:: ||:
Qy     216 VLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFDQV 275

Db     252 TLLKAAC-LDILILRICTRYTPEQDTMTFSDGLTLNRTQ-MHNAGFGPLTDLVFTFANQL 309
            ||   :: ::::::  :  |          . |.  | |:.  .  .    |.   . |:::: ::::
Qy     276 RLL-ESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSRF 334

Db     310 LPLEMDDTE 318
             | ::   |
Qy     335 RELKLQHKE 343

RESULT   14
ID   PCT-US92-02320A-4     STANDARD;     PRT;    462 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 4, Application PC/TUS9202320A.
XX
CC    Sequence 4, Application PC/TUS9202320A
CC     GENERAL INFORMATION:
CC       APPLICANT: Sloan-Kettering Institute, For Cancer Research
CC       TITLE OF INVENTION: METHODS FOR DETECTION AND TREATMENT OF CANCER
CC       NUMBER OF SEQUENCES: 6
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE: COOPER & DUNHAM
CC         STREET: 30 ROCKEFELLER PLAZA
CC         CITY: NEW YORK
CC         STATE: NEW YORK
CC         COUNTRY: U.S.A.
CC         ZIP: 10112
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE: Floppy disk
CC         COMPUTER: IBM PC compatible
CC         OPERATING SYSTEM: PC-DOS/MS-DOS
CC         SOFTWARE: PatentIn Release #1.24
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER: PCT US92 02320A
CC         FILING DATE:
```

```
CC         CLASSIFICATION:
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER: US 675,084
CC         FILING DATE: 22-MAR-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER: US 673,838
CC         FILING DATE: 22-MAR-1991
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME: WHITE, JOHN P
CC         REGISTRATION NUMBER: 28,678
CC         REFERENCE/DOCKET NUMBER: 38694-PCT
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE: (212) 977-9550
CC         TELEFAX: (212) 664-0525
CC         TELEX: (212) 422523 COOP UI
CC     INFORMATION FOR SEQ ID NO: 4:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH: 462 amino acids
CC         TYPE: amino acid
CC         TOPOLOGY: linear
CC       MOLECULE TYPE: protein
SQ   SEQUENCE   462 AA; 50771 MW; 1101259 CN;

  Query Match             12.6%;  Score 393;  DB 10;  Length 462;
  Best Local Similarity   29.7%;  Pred. No. 3.09e-25;
  Matches   74;  Conservative   60;  Mismatches  103;  Indels   12;  Gaps   12;

Db      88 CFVCQDKSSGYHYGVSACEGCKGFFRRSIQKNMVYTCHRDKNCIINKVTRNRCQYCRLQK 147
           | || |  :||||||| :||||||:||:||||  :   | |    :  | | |:|   |: || ||||:|
Qy      96 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 155

Db     148 CFEVGMSKESVRNDR-N-K-KKEVPKPECSESYTLTPEVGELIEKVRKAHQETF-PA-L 202
           |:|||| |  :  |  :|     :    ::: :   |            :     |      :||     ::: |  |
Qy     156 CYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 215

Db     203 C-QI-GKYTTNN-SSEQRVSL-DIDLWDKFSELSTKCIIKTVEFAKQLPGFTTLTIADQI 258
                 |        :     |       :: :  :     :: |: | ::  :  || :|||  ::: ||:
Qy     216 VLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFDQV 275

Db     259 TLLKAAC-LDILILRICTRYTPEQDTMTFSDGLTLNRTQ-MHNAGFGPLTDLVFAFANQL 316
            ||   :| :::|:: :  |          . |.  | |:| : :    |:   : |:::| ::::
Qy     276 RLL-ESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSRF 334

Db     317 LPLEMDDAE 325
             | ::   |
Qy     335 RELKLQHKE 343

RESULT   15
ID   PCT-US92-02320A-2     STANDARD;     PRT;    797 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 2, Application PC/TUS9202320A
XX
CC    Sequence 2, Application PC/TUS9202320A
CC     GENERAL INFORMATION:
CC       APPLICANT: Sloan Kettering Institute, For Cancer Research
CC       TITLE OF INVENTION: METHODS FOR DETECTION AND TREATMENT OF CANCER
CC       NUMBER OF SEQUENCES: 6
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE: COOPER & DUNHAM
CC         STREET: 30 ROCKEFELLER PLAZA
CC         CITY: NEW YORK
CC         STATE: NEW YORK
CC         COUNTRY: U.S.A.
CC         ZIP: 10112
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE: Floppy disk
CC         COMPUTER: IBM PC compatible
CC         OPERATING SYSTEM: PC-DOS/MS-DOS
CC        SOFTWARE:  PatentIn Release #1.24
CC       CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  PCT/US92/02320A
CC        FILING DATE:
CC        CLASSIFICATION:
CC       PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US 675,084
CC        FILING DATE:  22-MAR-1991
CC       PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US 673,838
CC        FILING DATE:  22-MAR-1991
CC       ATTORNEY/AGENT INFORMATION:
CC        NAME:  WHITE, JOHN P
CC        REGISTRATION NUMBER:  28,678
CC        REFERENCE/DOCKET NUMBER:  38694-PCT
CC       TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  (212) 977-9550
CC        TELEFAX:  (212) 644-0525
CC        TELEX:  (212) 422523 COOP UI
      INFORMATION FOR SEQ ID NO:  2:
         SEQUENCE CHARACTERISTICS:
          LENGTH:  797 amino acids
CC        TYPE:  amino acid
CC        TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   797 AA;  89281 MW;  3152972 CN;

  Query Match            12.6%;  Score 393;  DB 10;  Length 797;
  Best Local Similarity  29.7%;  Pred. No. 3.09e-25;
  Matches   74;  Conservative   60;  Mismatches  103;  Indels   12;  Gaps   12;

Db    423 CFVCQDKSSGYHYGVSACEGCKGFFRRSIQKNMVYICHRDKNCIINKVIRNRCQYCRLQK 482
          | || | :|| |||| :|||||:||:|||| :  | |   : | |:|  |: || |||:|
Qy     96 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 155

Db    483 CFEVGMSKESVRNDR-N-K-KKKEVPKPECSESYTLTPEVGELIEKVRKAHQETF-PA-L 537
          |:|||| | : | :|   :  ::: :  |       :   |    :||    ::: |  |
Qy    156 CYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPPVPELLLDALSPEQL 215

Db    538 C-QL-GKYTTNN-SSEQRVSL-DIDLWDKFSELSTKCIIKIVEFAKQLPGFITLTIADQI 593
             |      :   |    :: :  :   :: |: | ::  :  || :|||  |:: ||:
Qy    216 VLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFDQV 275

Db    594 TLLKAAC-LDILILRICTRYTPEQDTMTFSDGLTLNRTQ-MHNAGFGPLTDLVFAFANQL 651
           ||  :| :::|:: :  |       : |:  | |:| : :   |.  . |...| ....
Qy    276 RLL-ESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSRF 334

      652 LPLEMDDAE 660
            | ::  |
      335 RELKLQHKE 343

Search completed: Thu Dec 18 09:08:41 1997
Job time : 34 secs.
```

```
***********************************************************************

                     M P S R C H                              (TM)

***********************************************************************

          Release 2.1D John F. Collins, Biocomputing Research Unit
          Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
                Distribution rights by IntelliGenetics, Inc.

MPsrch_pp   protein - protein database search, using Smith-Waterman algorithm

   on             Thu Dec 18 08:56:29 1997   MasPar time 11.71 Seconds
                                             801.706 Million cell updates/sec
Tabular output not generated.

Title:             >US-08-826-361A-6
Description:       (1-416) from US08826361A.pep
Perfect Score:     3108
Sequence:          1 MNYSIPSNVTNLEGGPGRQT..........QQSMRLANLLMLLSHVRHAR 416

Scoring table.     PAM 150
                   Gap 11

Searched:          59021 seqs, 21210388 residues

Post-processing.   Minimum Match 0%
                   Listing first 45 summaries

Database:          swiss-prot34
                     1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
                     8:part8 9:part9 10:part10 11:part11

Statistics:        Mean 49.376;  Variance 92.220;  scale 0.601

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                SUMMARIES
                  %
 ult            Query
 No.   Score    Match Length DB  ID           Description              Pred. No.
-------------------------------------------------------------------------------
   1    1472     47.4    589  3  ESTR_CHICK   ESTROGEN RECEPTOR (ER   0.00e+00
   2    1473     47.4    595  3  ESTR_HUMAN   ESTROGEN RECEPTOR (ER   0.00e+00
   3    1455     46.8    599  3  ESTR_MOUSE   ESTROGEN RECEPTOR (ER   0.00e+00
   4    1448     46.6    600  3  ESTR_RAT     ESTROGEN RECEPTOR (ER   0.00e+00
   5    1381     44.4    575  3  ESTR_ORYLA   ESTROGEN RECEPTOR (ER   6.35e-303
   6    1327     42.7    574  3  ESTR_SALIR   ESTROGEN RECEPTOR (ER   2.46e-289
   7    1314     42.3    535  3  ESTR_SALSA   ESTROGEN RECEPTOR (ER   4.56e-286
   8     879     28.3    314  3  ESTR_OREAU   ESTROGEN RECEPTOR (ER   1.30e-177
   9     736     23.7    433  3  ERR2_HUMAN   STEROID HORMONE RECEP   1.61e-142
  10     725     23.3    521  3  ERR1_HUMAN   STEROID HORMONE RECEP   7.73e-140
  11     681     21.9    923  7  PRGR_MOUSE   PROGESTERONE RECEPTOR   3.86e-129
  12     666     21.4    933  7  PRGR_HUMAN   PROGESTERONE RECEPTOR   1.67e-125
  13     662     21.3    930  7  PRGR_RABIT   PROGESTERONE RECEPTOR   1.55e-124
  14     614     19.8    777  4  GCR_HUMAN    GLUCOCORTICOID RECEPT   5.96e-113
  15     608     19.6    776  4  GCR_XENLA    GLUCOCORTICOID RECEPT   1.65e-111
  16     597     19.2    786  7  PRGR_CHICK   PROGESTERONE RECEPTOR   7.26e-109
  17     586     18.9    783  4  GCR_MOUSE    GLUCOCORTICOID RECEPT   3.15e-106
  18     579     18.6    133  3  ESTR_BOVIN   ESTROGEN RECEPTOR (ER   1.50e-104
  19     577     18.6    771  4  GCR_CAVPO    GLUCOCORTICOID RECEPT   4.51e-104
  20     577     18.6    795  4  GCR_RAT      GLUCOCORTICOID RECEPT   4.51e-104
  21     562     18.1    981  6  MCR_RAT      MINERALOCORTICOID REC   1.74e-100
  22     557     17.9    899  1  ANDR_MOUSE   ANDROGEN RECEPTOR.      2.71e-99
  23     556     17.9    902  1  ANDR_RAT     ANDROGEN RECEPTOR.      4.69e-99
  24     554     17.8    709  1  ANDR_RABIT   ANDROGEN RECEPTOR (ER   1.41e-98
  25     553     17.8    919  1  ANDR_HUMAN   ANDROGEN RECEPTOR.      2.44e-98
  26     544     17.5    758  4  GCR_[illegible]  GLUCOCORTICOID RECEPT   3.33e-95
  27     525     16.9    984  6  MCR_HUMAN    MINERALOCORTICOID REC   3.30e-91
  28     517     16.6    111  3  ESTR_SHEEP   ESTROGEN RECEPTOR (ER   8.73e-90
  29     506     16.3    533  8  RRXB_HUMAN   RETINOIC ACID RECEPTO   3.49e-87
  30     505     16.2    458  8  RRXB_RAT     RETINOIC ACID RECEPTO   5.02e-87
  31     505     16.2    520  8  RRXB_MOUSE   RETINOIC ACID RECEPTO   5.02e-87
  32     501     16.1    597  6  NGFI_RAT     NERVE GROWTH FACTOR I   5.30e-86
  33     499     16.1    598  6  NGFI_CANFA   ORPHAN NUCLEAR RECEPT   1.57e-85
  34     493     15.9    462  8  RRXA_HUMAN   RETINOIC ACID RECEPTO   4.08e-84
  35     494     15.9    467  8  RRXA_RAT     RETINOIC ACID RECEPTO   2.37e-84
  36     494     15.9    467  8  RRXA_MOUSE   RETINOIC ACID RECEPTO   2.37e-84
  37     495     15.9    598  6  NGFI_HUMAN   EARLY RESPONSE PROTEI   1.38e-84
  38     487     15.7    439  8  RRXA_XENLA   RETINOIC ACID RECEPTO   1.06e-82
  39     489     15.7    601  6  NGFI_MOUSE   NUCLEAR HORMONE RECEP   3.57e-83
  40     473     15.2    598  8  RNR1_RAT     REGENERATING LIVER NU   2.06e-79
  41     471     15.2    598  6  NOT_HUMAN    IMMEDIATE-EARLY RESPO   6.06e-79
  42     471     15.2    595  7  NOT1_MOUSE   PROBABLE NUCLEAR HORM   6.06e-79
  43     469     15.1    467  8  RRXA_CHICK   RETINOIC ACID RECEPTO   1.79e-78
  44     464     14.9    463  8  RRXG_HUMAN   RETINOIC ACID RECEPTO   2.65e-77
  45     464     14.9    463  8  RRXG_MOUSE   RETINOIC ACID RECEPTO   2.65e-77

                                ALIGNMENTS

RESULT    1
ID   ESTR_CHICK     STANDARD;      PRT;   589 AA.
AC   P06212;
DT   01-JAN-1988 (REL. 06, CREATED)
DT   01-JAN-1988 (REL. 06, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   ESTROGEN RECEPTOR (ER).
GN   ESR.
OS   GALLUS GALLUS (CHICKEN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;
OC   GALLIFORMES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 86247578.
RA   KRUST A., GREEN S., ARGOS P., KUMAR V., WALTER P., BORNERT J.-M.,
RA   CHAMBON P.;
RL   EMBO J. 5:891-897(1986).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- IN THE ABSENCE OF LIGAND, STEROID HORMONE RECEPTORS ARE THOUGHT
CC       TO BE WEAKLY ASSOCIATED WITH NUCLEAR COMPONENTS; HORMONE BINDING
CC       GREATLY INCREASES RECEPTOR AFFINITY. THE HORMONE-RECEPTOR
CC       COMPLEX APPEARS TO RECOGNIZE DISCRETE DNA SEQUENCES UPSTREAM OF
CC       TRANSCRIPTIONAL START SITES.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; X03805; G63380; -.
DR   PIR; S07192; S07192.
DR   HSSP; P06536; 1GDC.
DR   PROSITE; PS00031; STEROID_FINGER.
DR   TRANSFAC; T00264; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
FT   DOMAIN        1    178       MODULATING.
FT   DNA_BIND    179    244       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     179    199       C4-TYPE.
FT   ZN_FING     215    239       C4-TYPE.
FT   DOMAIN      245    304       HINGE.
FT   DOMAIN      305    589       STEROID-BINDING.
SQ   SEQUENCE   589 AA;  66740 MW;  85FF1F4A CRC32;
```

```
  Query Match             47.4%;  Score 1472;  DB 3;  Length 589;
  Best Local Similarity   51.8%;  Pred. No. 0.00e+00;
  Matches  212;  Conservative   95;  Mismatches   85;  Indels   18;  Gaps   11;

Db      100 spvvflqtapqlspflhhhsqqvpyylenepqsfqmreaappafyrpssdnrrhslrerm 159
            || |:   |: :|||:: |: |       | | : ::       :    ::   : :  :: :
Qy       22 SPNVLWPTPGHLSPLVVHR-QLSHLYAEPQKSPWCEAPSLEHTL-PVNRETLKRKVSGNR 79

Db      160 sstnekqslsmestketrycavcndyasgyhygvwscegckaffkrsiqghndymcpatn 219
              ::    |: :    | :::::|||||:|||||||||||||||||||||||||||:|||||
Qy       80 CASPVTGP-G -SKEDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAIN 135

Db      220 qctidknrrksqqacrlrkcyevgmmkggirkdrrggqmmkqkrqreeqdsrngeasste 279
            |||||||||||||||||||||||||||:| | |::| |   :::: |   '||     | |   .
Qy      137 QCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQRSALEQLHCAGKAKRSG 196

Db      280 lraptlwtspl--khnkknspalsltaeqmvsalleaeppivysrydpnrpfneasmmtl 339
              :|| ::   |::        :   |||   :||:| :|||||||  |       |: || |||||
Qy      197 GHAPRVR E LLL ----D--ALS--PFQLVLTLLEAPPHVLIS-RPSAPFTEASMMMS 244

Db      340 ltnladrelvhminwakrvpgfvdltlhdqvhllecawleilmiglvwrsmehpgkllfa 399
            ||:|||:||||||:|||::||||:':| ||||:|||   |:|:||:||:|||::||||||:||
Qy      245 LTKLADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFA 304

Db      400 pnllldrnqgkcveqmveifdmllataarfrmmnlqgeefvclksiillnsgvytflsst 459
            |:|:|||::||||||||:||||||||||||::||| ::||  |::|:|::|||||::| ::::|
Qy      305 PDLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVIAT 364

Db      460 lksleerdyihrvldkitdtlihlmaksglslqqqhrrlaqllllilshirh 510
              :    |    | :|: :||:|: ::||||:| |||   ||| ||::|||:||
Qy      365 QDADSSRKLAH-LLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRH 414

RESULT      2
ID   ESTR_HUMAN      STANDARD;      PRT;   595 AA.
AC   P03372;
DT   21-JUL-1986 (REL. 01, CREATED)
DT   01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DT   01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE   ESTROGEN RECEPTOR (ER).
GN   ESR.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 86122927.
RA   GREENE G.L., GILNA P., WATERFIELD M., BAKER A., HORT Y., SHINE J.;
RL   SCIENCE 231:1150-1154(1986).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 86146892.
RA   GREEN S., WALTER P., KUMAR V., KRUST A., BORNERT J.-M., ARGOS P.,
RA   CHAMBON P.;
RL   NATURE 320:134-139(1986).
RN   [3]
RP   STRUCTURE BY NMR OF 180-262
RX   MEDLINE; 91061891.
RA   SCHWABE J.W.E., NEUHAUS D., RHODES D.;
RL   NATURE 348:458-461(1990).
RN   [4]
RP   VARIANT VAL-400.
RX   MEDLINE; 90005402.
RA   TORA L., MULLICK A., METZGER D., PONGLIKITMONGKOL M., PARK I.,
RA   CHAMBON P.;
RL   EMBO J. 8:1981-1986(1989).
RN   [5]
RP   VARIANT ALA-447.
RX   MEDLINE; 92250650.
RA   REESE J.C., KATZENELLENBOGEN B.S.;
RL   J. BIOL. CHEM. 267:9868-9873(1992).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- IN THE ABSENCE OF LIGAND, STEROID HORMONE RECEPTORS ARE THOUGHT
CC       TO BE WEAKLY ASSOCIATED WITH NUCLEAR COMPONENTS; HORMONE BINDING
CC       GREATLY INCREASES RECEPTOR AFFINITY. THE HORMONE-RECEPTOR
CC       COMPLEX APPEARS TO RECOGNIZE DISCRETE DNA SEQUENCES UPSTREAM OF
CC       TRANSCRIPTIONAL START SITES.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; X03635; G31234; -.
DR   EMBL; M12674; G182193; -.
DR   PIR; A03244; QRHUE.
DR   HSSP; P06536; 1GDC.
DR   MIM; 133430; -.
DR   PROSITE; PS00031; STEROID_FINGER.
DR   TRANSFAC; T00261; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING; POLYMORPHISM.
FT   DOMAIN        1    184       MODULATING.
FT   DNA_BIND    185    250       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     185    205       C4-TYPE.
FT   ZN_FING     221    245       C4-TYPE.
FT   DOMAIN      251    310       HINGE.
FT   DOMAIN      311    551       STEROID-BINDING.
FT   VARIANT     400    400       G -> V (DESTABILIZES THE RECEPTOR AND
FT                                DECREASES ITS AFFINITY FOR ESTRADIOL AT
FT                                25 DEGREES CELSIUS, BUT NOT AT 4 DEGREES
FT                                CELSIUS).
FT   VARIANT     447    447       C -> A (LOSS OF HORMONE BINDING CAPACITY
FT                                AND TEMPERATURE-SENSITIVE LOSS IN DNA-
FT                                BINDING).
SQ   SEQUENCE   595 AA;  66216 MW;  D6945BB9 CRC32;

  Query Match             47.4%;  Score 1473;  DB 3;  Length 595;
  Best Local Similarity   57.6%;  Pred. No. 0.00e+00;
  Matches  196;  Conservative   84;  Mismatches   45;  Indels   15;  Gaps   11;

Db      178 saketrycavcndyasgyhygvwscegckaffkrsiqghndymcpatnqctidknrrksc 237
            | :::::||||:|||||||||||||||||||||||||||||||||:|||||||||||||||||
Qy       89 SKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSC 148

Db      238 qacrlrkcyevgmmkggirkdrrggrmlkhkrqrddgegrgevgsagdmraanlwpsplm 297
            |||||||||||||:| | |::| | |:::  |||    | :    :: : :|::     :|
Qy      149 QACRLRKCYEVGMVKCGSRRERCGYRLVR--RQRSADE-QLHCAGKA-KRSGG  HAP-R 201

Db      298 ikrskknslalsltadqmvsalldaeppilyseydptrpfseasmmglltnladrelvh 356
             : |      |   :|::::|:| :||:|||| :| |    |: ||:|||||    ||:|||:||||
Qy      202 V-RELL--LD-ALSPEQLVLILLEAEPPHVLISR--PSAPFTEASMMMSLTKLADKELVH 255

Db      357 minwakrvpgfvdltlhdqvhllecawleilmiglvwrsmehpgkllfapnllldrnqgk 416
            ||.|||..||||.|.| |||.|||  |.|.||.||:|||..|||||.|||.|.|||..||
Qy      256 MISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGK 315

Db      417 cvegmveifdmllatssrfrmmnlqgeefvclksiillnsgvytflsstlksleekdhih 476
            ||||::||||||||||:||||| ::||  |::|:|::|||||::| ::::|   :  :     :
Qy      316 CVEGILEIFDMLLATISRFRELKLQHKEYLCVKAMILLNSSMYPLVIATQDA-DSSRKLA 374

Db      477 rvldkitdtlihlmakagltlqqqhqrlaqllllilshirh 516
            ::|: :||:|: ::||:|:: |||  ||| ||::|||:||
Qy      375 HLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRH 414

RESULT      3
ID   ESTR_MOUSE      STANDARD;      PRT;   599 AA.
AC   P19785;
DT   01-FEB-1991 (REL. 17, CREATED)
DT   01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
DT   01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE   ESTROGEN RECEPTOR (ER).
GN   ESR OR ESTR.
```

```
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=UTERUS;
RX   MEDLINE; 91042558.
RA   WHITE R., LEES J.A., NEEDHAM M., HAM J., PARKER M.;
RL   MOL. ENDOCRINOL. 1:735-744(1987).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- IN THE ABSENCE OF LIGAND, STEROID HORMONE RECEPTORS ARE THOUGHT
CC       TO BE WEAKLY ASSOCIATED WITH NUCLEAR COMPONENTS; HORMONE BINDING
CC       GREATLY INCREASES RECEPTOR AFFINITY. THE HORMONE-RECEPTOR
CC       COMPLEX APPEARS TO RECOGNIZE DISCRETE DNA SEQUENCES UPSTREAM OF
CC       TRANSCRIPTIONAL START SITES.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; M38651; G193180; -.
DR   PIR; A40061; A40061.
DR   HSSP; P06536; 1GDC.
DR   PROSITE; PS00031; STEROID_FINGER.
DR   TRANSFAC; T00259; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
FT   DOMAIN        1    188       MODULATING.
FT   DNA_BIND    189    254       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     189    209       C4-TYPE.
FT   ZN_FING     225    249       C4-TYPE.
FT   DOMAIN      255    599       STEROID-BINDING.
FT   DOMAIN       64     72       POLY-ALA.
SQ   SEQUENCE   599 AA;  66455 MW;  FED9EA18 CRC32;

  Query Match             46.8%,  Score 1455;  DB 3;  Length 599;
  Best Local Similarity   51.7%;  Pred. No. 0.00e+00;
  Matches  214;  Conservative   93;  Mismatches   87;  Indels   20;  Gaps   14;

Db     109 spsplmllhppqlspflqphgqqvpyylenepsay-avrdtgppafyrsnsdnrrqngr 166
           : || :|     | :||| |  | |       | | . |     |          .  |  |
Qy      20 TTSPNVLWPTPGHLSP-LVVHRQLSHLYAEPQKSPWCEARSLEHTLPVNRETLKRKVSG- 77

Db     167 erlsssnekgnmimesaketrycavcndyasgyhygvwscegckaffkrsiqghndymcp 226
           :|  :|   |       | :::::||||:||||||||||||||||||||||||||||||:||
Qy      78 NR-CASPVTGP---GSKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICP 133

Db     227 atnqctidknrrkscqacrlrkcyevgmmkggirkdrrggrmlkhkrqrddlegrnemga 286
           ||||||||||||||||||||||||||||||:| | |::| | |:::  |||    : :    |
Qy     134 ATNQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVR--RQRSA-DEQLH-CA 189

Db     287 sgdmraanlwpsplvikhtkknspalsltadqmvsalldaeppmiyseydpsrpfseasm 346
           :  :|::    :| | :     :  |||  ::|:| .||:|||| :      || ||.||||
Qy     190 GKAKRSGG--HAPRV-RELLLD--ALS--PEQLVLTLLEAEPPHVLIS-RPSAPFTEASM 241

Db     347 mglltnladrelvhminwakrvpgfgdlnlhdqvhllecawleilmiglvwrsmehpgkl 406
           |  ||:|||:|||||:|||::|||  :|:| |||:|||  |:|:||:|:|:|||  |||||
Qy     242 MMSLTKLADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKL 301

Db     407 lfapnllldrnqgkcvegmveifdmllatssrfrmmnlqgeefvclksiillnsgvytfl 466
           :|||:|:|||::||||||::|||||||||||:||||  :.||   |  |  .||||||:| ..
Qy     302 IFAPDLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLV 361

Db     467 sstlksleekdhihrvldkitdtlihlmakagltlqqqhrrlaqllllilshirh 520
           :.|  .  .     . ..|. .||.|. ..||. . . |||   |||  ||..||||.||
Qy     362 TATQDA-DSSRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRH 414


RESULT     4
ID   ESTR_RAT       STANDARD;      PRT;   600 AA.
AC   P06211;
DT   01-JAN-1988 (REL. 06, CREATED)
DT   01-JAN-1988 (REL. 06, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   ESTROGEN RECEPTOR (ER).
GN   ESR OR ESTR.
OS   RATTUS NORVEGICUS (RAT).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WISTAR;
RA   MURAMATSU M.;
RL   SUBMITTED (MAR-1987) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 87174780.
RA   KOIKE S., SAKAI M.;
RL   NUCLEIC ACIDS RES. 15:2499-2513(1987).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=UTERUS;
RA   MAGGI A M.A.;
RL   SUBMITTED (JUN-1991) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- IN THE ABSENCE OF LIGAND, STEROID HORMONE RECEPTORS ARE THOUGHT
CC       TO BE WEAKLY ASSOCIATED WITH NUCLEAR COMPONENTS; HORMONE BINDING
CC       GREATLY INCREASES RECEPTOR AFFINITY. THE HORMONE-RECEPTOR
CC       COMPLEX APPEARS TO RECOGNIZE DISCRETE DNA SEQUENCES UPSTREAM OF
CC       TRANSCRIPTIONAL START SITES.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; Y00102; G56111; -.
DR   EMBL; X61098; G56121; -.
DR   PIR; S07379; S07379.
DR   HSSP; P06536; 1GDC.
DR   PROSITE; PS00031; STEROID_FINGER.
DR   TRANSFAC; T00258; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
FT   DOMAIN        1    189       MODULATING.
FT   DNA_BIND    190    255       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     190    210       C4-TYPE.
FT   ZN_FING     226    250       C4-TYPE.
FT   DOMAIN      256    600       STEROID-BINDING.
FT   DOMAIN       64     71       POLY-ALA.
FT   CONFLICT    488    488       N -> T (IN REF. 3).
SQ   SEQUENCE   600 AA;  67030 MW;  4CCC59B2 CRC32;

  Query Match             46.6%,  Score 1448;  DB 3;  Length 600;
  Best Local Similarity   51.0%;  Pred. No. 0.00e+00;
  Matches  211;  Conservative   97;  Mismatches   86;  Indels   20;  Gaps   14;

Db     109 spsplmllhpphvspflqphgqqvpyylenepsay-avrdtgppafyrsnsdnrrqnsr 167
           . || .|     | |.:| |  | .       | | . |:      |          :: :|: :|
Qy      20 TTSPNVLWPTPGHLSP-LVVHRQLSHLYAEPQKSPWCEARSLEHTLPVNRETLKRKVSG- 77

Db     168 erlssssekgnmimesaketrycavcndyasgyhygvwscegckaffkrsiqghndymcp 227
           .|  .|.  |      | .....||||:||||||||||||||||||||||||||||||:||
Qy      78 NR-CASPVTGP---GSKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICP 133

Db     228 atnqctidknrrkscqacrlrkcyevgmmkggirkdrrggrmlkhkrqrddlegrnemgt 287
           |||:||||||||||||||||||||||||||:| | |::| : |:::: |  |: : :    |
Qy     134 ATNQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQRSADE-QLHCA-GK 191

Db     298 sgdmraanlwpsplvikhtkknspalsltadqmvsalldaeppliyseydpsrpfseasm 347
           :  :|::    :| | :     :  |||  ::|:| :||:|||| :      |! ||:||||
Qy     192 A--KRSGG--HAPRV-RELLLD--ALS--PEQLVLTLLEAEPPHVLIS-RPSAPFTEASM 241

Db     348 mqlltnladrelvhminwakrvpqfqdlnlhdqvhllecawleilmiglvwrsmehpgkl 407
           |  ||:|||: |||||:|||::||| :|:| |||:|||  |:|:||:||:|||::|||||
Qy     242 MMSLFKLADKLVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKL 301

Db     408 lfapllldrnqgkcveqmveifdmllatssrfrmmnlqqeefvclksiillnsgvytfl 467
           :|||:|:|||::||||||::||||||||||:||||  ::||  |::|:|::||||||::| ::
Qy     302 IFAPDLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLV 361

Db     468 sstlksleekdhihrvldkindtlihlmakaqltlqqhrrlaqlllilshirh 521
           ::|   :  :      :  ::|:  :  |:|:  ::||:|::  |||   |||  ||::|||:||
Qy     362 TATQDA-DSSRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRH 414

RESULT     5
ID   ESTR_ORYLA     STANDARD;      PRT;   575 AA.
AC   P50241;
DT   01-OCT-1996 (REL. 34, CREATED)
DT   01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
     01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
     ESTROGEN RECEPTOR (ER).
GN   ESR.
OS   ORYZIAS LATIPES (MEDAKA FISH).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; PISCES; GNATHOSTOMATA;
OC   OSTEICHTHYES; ACTINOPTERYGII; ATHERINIFORMES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RA   OKADA H.;
RL   SUBMITTED (APR-1994) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; D28954; G602326; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
FT   DOMAIN        1    140       MODULATING.
FT   DNA_BIND    141    206       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     141    161       C4-TYPE.
FT   ZN_FING     177    201       C4-TYPE.
FT   DOMAIN      207    575       STEROID-BINDING.
     DOMAIN      254    262       POLY-GLY.
     DOMAIN      275    278       POLY-LEU.
     DOMAIN      466    470       POLY-LEU.
FT   DOMAIN      531    534       POLY-GLY.
SQ   SEQUENCE   575 AA;  62987 MW;  D9FE40C0 CRC32;

  Query Match             44.4%;  Score 1381;  DB 3;  Length 575;
  Best Local Similarity   52.1%;  Pred. No. 6.35e-303;
  Matches  223;  Conservative   83;  Mismatches   93;  Indels   29;  Gaps   16;

Db      52 ppseqslqs-lqsaptsplvfvpssprlspfmhpp-shhylettstpvyrsshqqasre 108
           |::  :|::  |    |||  |: |::  :|||:: |    || | |    :|    :    :
Qy       6 PSNVTNLEGGPGRQTTSPNVLWPIPGHLSPLVVHRQLSHLYAEPQKSPWCEARSLEHTLP 65

Db     109 dqeqsredtesiqelqagagaqgfemakdtrfcavcsdyasgyhygvwscegckaffkrs 168
                :      | |:  |::  :|     :  :|::|||||||||||||||||||||||||||
Qy      66 VNRETLKRKVS-GNRCASPVTGPNSK-RDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRS 123

Db     169 iqghndymcpatnqctidrnrrksc-qacrlrkcyevgmvkggvrkdririlrrdkrrtqv 228
           |||||||:|||||||||:|||||||||||||||||||||:|  | |::|      |     ||
Qy     124 IQGHNDYICPATNQCIIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGY-RIV RR--- 178

Db     229 adqlkvvkrpnkkhyhydirkrs-tqqqpqqqlsvtslppeqvllllqgaeppllc 287
             ::    ::|:  :|   |: : | ||  :          | :  :::|||::|  |   ||||  :|
Qy     179    QF   FAAEL -LHCAGKAKRS GGHAPRVRELL-LEALSPEQLVLTLLEAEPPHVLI 229
```

```
Db     288 srqklsrpytevtmmtlltsradkelvhmiawakklpqflqlslhdqvlllesswlevlm 347
           ||    : |:|| :||  || :|||||||||:||||:|||::|||  |||  ||||  |:||||
Qy     230 SRPS-A-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLM 287

Db     348 igliwrsihcpgkllfaqdllldrnegdcveqmteifdmllatasrfrvlklkpeefvcl 407
           :||:||||  |||||||:||:||||:||:  ||||:  ||||||||||:|||| |||   |:::
Qy     288 MGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCV 347

Db     408 kaillnsgafsfctqtmeplhnsaavqsmldtitdalihyisqsqylaqeqarrqaqll 467
           ||:||||||: ::: |:| ::   :|   :   :|:::||||:   ': ||   :|:|: | | ||
Qy     348 KAMILLNSSMYPLVTATQDA-DSSRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLL 406

Db     468 lllshirh 475
           :||||:||
Qy     407 MLLSHVRH 414

RESULT     6
ID   ESTR_SALIR     STANDARD;      PRT;   574 AA.
AC   P16058;
DT   01-APR-1990 (REL. 14, CREATED)
DT   01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   ESTROGEN RECEPTOR (ER).
GN   ESR.
OS   SALMO IRIDEUS (RAINBOW TROUT).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; PISCES; GNATHOSTOMATA;
OC   OSTEICHTHYES; ACTINOPTERYGII; SALMONIFORMES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 91006824.
RA   PAKDEL F., LE GAC F., LE GOFF P., VALOTAIRE Y.;
RL   MOL. CELL. ENDOCRINOL. 71:195-204(1990).
RN   [2]
RP   SEQUENCE OF 150-574 FROM N.A.
RC   TISSUE=LIVER;
RX   MEDLINE; 89127284.
RA   PAKDEL F., LE GUELLEC C., VAILLANT C., LE ROUX M.G., VALOTAIRE Y.;
RL   MOL. ENDOCRINOL. 3:44-51(1989).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- IN THE ABSENCE OF LIGAND, STEROID HORMONE RECEPTORS ARE THOUGHT
CC       TO BE WEAKLY ASSOCIATED WITH NUCLEAR COMPONENTS; HORMONE BINDING
CC       GREATLY INCREASES RECEPTOR AFFINITY. THE HORMONE-RECEPTOR
CC       COMPLEX APPEARS TO RECOGNIZE DISCRETE DNA SEQUENCES UPSTREAM OF
CC       TRANSCRIPTIONAL START SITES.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; M31559; G213784; -.
DR   PIR; A37197; A37197.
DR   PIR; A40070; A40070.
DR   HSSP; P06536; 1GDC.
DR   PROSITE; PS00031; STEROID_FINGER.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
FT   DOMAIN        1    146       MODULATING (POTENTIAL).
FT   DNA_BIND    147    212       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     147    167       C4-TYPE.
FT   ZN_FING     183    207       C4-TYPE.
FT   DOMAIN      213    574       STEROID-BINDING (POTENTIAL).
SQ   SEQUENCE   574 AA;  62971 MW;  B391A0C7 CRC32;

  Query Match             42.7%;  Score 1327;  DB 3;  Length 574;
  Best Local Similarity   57.5%;  Pred. No. 2.46e-289;
  Matches  195;  Conservative   64;  Mismatches   61;  Indels   20;  Gaps   15;

Db     142 etrycavcsdfasgyhygvwscegckaffkrsiqghndymcpatnqctidrnrrksacqac 202
```

```
              ::::||||||:||||||||||||||||||||||||||||:||||||||:|:|||||||||
Qy       92 DAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQAC 151

Db      203 rlrkcyevgmvkqglrkdrgg-rvlrkdkrycgpagdrekptvtwstgqrpqdggrnsss 261
            ||||||||||||| | |::| | |::|:: |    :| | |     : : : |: |': ::
Qy      152 RLRKCYEVGMVKCGSRRERCGYRLVRRQ-R---SA-D-EQLHCA-GKAKRS--GGH--AP 200

Db      262 slnggggwrgpritmppeqvlfllqggtp-alcsrqkvarpytevtmmtlltsmadkelv 320
              :         : :: | : || ||::: |   | ||    | |:|| :||   || :||||||
Qy      201 PVREL--LLDA-LS-PEQLVLTLLEAEPPHVLISRPS-A-PFTEASMMMSLTKLADKELV 254

Db      321 hmiawakkvpgfqelslhdqvqllesswl-vlmigliwrslhcpgklifaqdllldrs-g 380
            |||:||||:|:| |||| |||:|||| |:||||:||:||||  |||||| ||:||| ||
Qy      255 HMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEG 314

Db      381 dcvegmaeifdmllatvsrfgmlklkpeefvclkaiillnpgafsfcsnsveslhnssav 440
             ||||: ||| |:|||| |||  |||    |::|:||:||||:: ::: : : ::  :|  :
Qy      315 KCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTAIQDA-DSSRK 373

Db      441 esmldnitdalihhishsgasvqqqprrqaqllllshirh 481
             :|: .||||.  |. || | |||. | | ||:|||||:||
Qy      374 AHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRH 414


RESULT     7
ID   ESTR_SALSA     STANDARD;      PRT;   535 AA.
AC   P50242;
DT   01-OCT-1996 (REL. 34, CREATED)
DT   01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   ESTROGEN RECEPTOR (ER) (FRAGMENT)
GN   ESR.
OS   SALMO SALAR (ATLANTIC SALMON).
OC   EUKARYOTA, METAZOA, CHORDATA, VERTEBRATA, PISCES; GNATHOSTOMATA.
OC   OSTEICHTHYES; ACTINOPTERYGII; SALMONIFORMES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RA   ROGERS S.A., LLEWELLYN L., RAMSURN V.P., SWEENEY G.E., WIGHAM T.,
RL   SUBMITTED (JUL-1995) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY)
     -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
         NUCLEAR HORMONE RECEPTORS
     EMBL; X89959; G929902; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING
FT   NON_TER       1      1
FT   DOMAIN       <1    104       MODULATING (POTENTIAL).
FT   DNA_BIND    105    170       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     105    125       C4-TYPE.
FT   ZN_FING     141    165       C4-TYPE.
FT   DOMAIN      171    535       STEROID-BINDING (POTENTIAL)
SQ   SEQUENCE   535 AA;  58900 MW;  0F71CD0C CRC32;

  Query Match              42.3%;  Score 1314;  DB 3;  Length 535;
  Best Local Similarity   56.9%;  Pred. No. 4.55e-286;
  Matches   195;  Conservative   61;  Mismatches   66;  Indels   21;  Gaps   14;

Db      101 etrycavcsdfasgyhygfwscegckaffkrsiqghndymcpatnqctmdrnrrkscqac 160
            ::::|||||||:||||||| ||||||||||||||||||||||:||||||||:|:||||||||
Qy       92 DAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQAC 151

Db      161 rlrkcyevgmvkqglrkdrgg-rvlrkdkrycgpagdrekpygdlehrtappqdggrnss 219
            ||||||||||||| | |::| | |::|:: |    :| :  .     .      . |  : ||
Qy      152 RLRKCYEVGMVKCGSRRERCGYRLVRRQ-R---SADE-QLHCAG---K-AK-RSGGHAPR 201

Db      220 ssslsggggwrgpritmppeqvlfllqgaeppalcsrqkvarpytevtmmtlltsmadke 279
```

```
              |            : :: | : || ||:: |   | ||    | |:|| :||   || :||||
Qy      202 VRELL---L-DA-LS-PEQLVLTLLEAEPPHVLISRPS-A-PFTEASMMMSLTKLADKE 252

Db      280 lvhmiawakkvpgfqelslhdqvqllesswlevlmigliwrslhcpgklifaqdllldrs 339
            |||||:||||:|:|| |||| |||:|||| |:||||:||:||||  ||||||! ||:|||
Qy      253 LVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRD 312

Db      340 egdcvegmaeifdmllatvsrfgmlklkpeefvclkaiillnsgafsfcsnsveslhnss 399
            || ||||: ||||||||| |||| |||   |::|:||:||||||: ::: : : ::  :|
Qy      313 EGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTAIQDA-DSSR 371

Db      400 avesmldnitdalihhishsgasvqqqprrqvqllllshirh 442
             .  .|. .||||:  |. || | |||: |   ||:||||:||
Qy      372 KLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRH 414


RESULT     8
ID   ESTR_OREAU     STANDARD;      PRT;   314 AA.
AC   P50240;
DT   01-OCT-1996 (REL. 34, CREATED)
DT   01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   ESTROGEN RECEPTOR (ER) (FRAGMENT).
GN   ESR.
OS   OREOCHROMIS AUREUS.
OC   EUKARYOTA, METAZOA, CHORDATA, VERTEBRATA, PISCES, GNATHOSTOMATA;
OC   OSTEICHTHYES; ACTINOPTERYGII; PERCIFORMES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BLOOD;
RX   MEDLINE; 96272302.
RA   TAN N., LAM T., DING J.;
RL   DNA SEQ. 5:359-370(1995).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; Z46665; E212276; -.
DR   EMBL; Z46666; E212276; JOINED.
DR   EMBL; Z46667; E212276; JOINED.
DR   EMBL; Z46668; E212276; JOINED.
DR   EMBL; Z46669; E212276; JOINED.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
FT   NON_TER       1      1
FT   DOMAIN       <1    314       STEROID-BINDING (POTENTIAL).
SQ   SEQUENCE   314 AA;  34871 MW;  1B1E3899 CRC32;

  Query Match              28.3%;  Score 879;  DB 3;  Length 314;
  Best Local Similarity   58.4%;  Pred. No. 1.30e-177;
  Matches   118;  Conservative   40;  Mismatches   40;  Indels    4;  Gaps    4;

Db        1 vllllqgaeppilcsrqkmsrpytevtimtlltsmadkelvhmitwakklpgflqlslhd 60
            || ||::  | :| |:    : |:|| ::'  || :|||||||'||:||||:|'|:::||  |
Qy      216 VLTLLEAEPPHVLISRPS-A-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 273

Db       61 qvlllesswlevlmigliwrslqcpgklifeedllldrnegtcvegmaeifdmllatv-r 119
            || |||| |.||||.||.||||. ||||||   ||.|'|:|| ||||: ||||'||||   |
Qy      274 QVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSR 333

Db      120 frvlklkpeefvclkaiillnsgafsfctgtmeplhdsvavqhmldtitdalifhishfg 179
            || |||   |..|.||.||||. ... |.| ..    |   . |.|...||||:  |:  |
Qy      334 FRELKLQHKEYLCVKAMILLNSSMYPLVIATQDA-DSSRKLAHLLNAVIDALVWVIAKSG 392

Db      180 csaqqqsrrqaqllllshirh 201
             |:'||| | | ||:||||:'|
Qy      393 ISSQQQSMRLANLLMLLSHVRH 414
```

```
RESULT     9
ID   ERR2_HUMAN     STANDARD;      PRT;    433 AA
AC   P11475;
DT   01-OCT-1989 (REL. 12, CREATED)
DT   01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)
DT   01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE   STEROID HORMONE RECEPTOR ERR2.
GN   ERR2.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=HEART;
RX   MEDLINE; 88122546.
RA   GIGUERE V., YANG N., SEGUI P., EVANS R.M.;
RL   NATURE 331:91-94(1988).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
         NUCLEAR HORMONE RECEPTORS.
     EMBL; X51417; G36611; -.
DR   PIR; B29345; B29345.
DR   HSSP; P06536; 1GDC.
DR   PROSITE; PS00031; STEROID_FINGER.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER.
FT   DNA_BIND    103    168       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     103    123       C4-TYPE.
FT   ZN_FING     139    163       C4-TYPE.
SQ   SEQUENCE   433 AA;  48287 MW;  673BF88F CRC32;

  Query Match             23.7%;  Score 736;  DB 3;  Length 433;
  Best Local Similarity  43.6%;  Pred. No. 1.61e-142;
  Matches  116;  Conservative   55;  Mismatches  84;  Indels  11;  Gaps   9;

Db     101 rlclvcgdiasgyhygvasceackaffkrtiqgnieyscpatneceitkrrrkscqacrf 160
           ::| ||:| ||||||||| |||:||||||||:|||: :| ||||||:| | | |||||||||:
Qy      94 HFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRL 153

Db     161 mkclkvgmlkegvrldrv--r-ggrqky-krrldsenspylslqisppakkpltkivs-- 214
             ||  |||:| | | :|    |   ||:      :|       :  |   :|  :  |   :|
Qy     154 RKCYEVGMVKGGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPE 213

Db     215 yl-l-vaepdklyam-ppddvpegdikalttlcdladrelvfliswakhipgfsnltlgd 271
             | | : |::    : :    |   :     : :| |||:||| :||||| |||| :|:| |
Qy     214 QLVLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 273

       272 qmsllqsawmeililgivyrslpyddklayaedyimdeehsrlv-gllelyrailqlvrr 330
           |: ||:| |||:|::|:: ||:     || :| | ::| : :: | |:||:: :|     |
       274 QVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSR 333

Db     331 ykklkvekeefvmlkalalansdsmy 356
           :: ||::  |:: :||: | || ||
Qy     334 FRELKLQHKEYLCVKAMILLNS-SMY 358


RESULT     10
ID   ERR1_HUMAN     STANDARD;      PRT;    521 AA.
AC   P11474;
DT   01-OCT-1989 (REL. 12, CREATED)
DT   01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)
DT   01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE   STEROID HORMONE RECEPTOR ERR1.
GN   ERR1.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=KIDNEY;
RX   MEDLINE; 88122546.
RA   GIGUERE V., YANG N., SEGUI P., EVANS R.M.;
RL   NATURE 331:91-94(1988).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; X51416; G36609; -.
DR   PIR; A29345; A29345.
DR   HSSP; P06536; 1GDC.
DR   PROSITE; PS00031; STEROID_FINGER.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER.
FT   DNA_BIND    176    241       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     176    196       C4-TYPE.
FT   ZN_FING     212    236       C4-TYPE.
SQ   SEQUENCE   521 AA;  55742 MW;  E2D723EA CRC32;

  Query Match             23.3%;  Score 725;  DB 3;  Length 521;
  Best Local Similarity  40.1%;  Pred. No. 7.73e-140;
  Matches  130;  Conservative   67;  Mismatches 117;  Indels  10;  Gaps   7;

Db     174 rlclvcgdvasgyhygvasceackaffkrtiqgsieyscpasneceitkrrrkacqacrf 233
           ::| ||:| ||||||||| |||:||||||||:|||  :| |||:|:| | | |||:|||||:
Qy      94 HFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRL 153

Db     234 tkclrvgmlkegvrldrvrggrqkykrrpevdplpfpg-p-fpagplavaggprktaapv 291
             ||  |||:| | | :|  | |    ::|: : |  :| :  ::|   :           | :
Qy     154 RKCYEVGMVKGGSRRERC-GYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSP 212

Db     292 nalvshllvvepeklyampdpagpdghlpavatlcdlfdreivvtiswaksipgfsslsl 351
           : ||  ||  ||  :  :: |::|     : : :|  | |:|:|  ||||| |||| |||
Qy     213 EQLVLTLLEAEPPHVL-ISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSL 271

Db     352 sdqmsvlqsvwmevlvlgvaqrslplqdelafaedlvld-eegaraaglgelgaallqlv 410
             ||: :|:| ||||||::|: ||:      | || ||||| :||     |: |:   ||
Qy     272 FDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATT 331

Db     411 rrlqalrlereeyvllkalalansdsvhiedeprlwssceklllh--eal---leyeagra 465
             |:: |:|:: ||: :||: | ||     :     :    |  || |   :|:    |    :::
Qy     332 SRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKS 391

Db     466 gpgggaerrragrllltlpllrqt 489
           | ::  :  | : ||: |: :|::
Qy     392 GISSQQQSMRLANLLMLLSHVRHA 415


RESULT   11
ID   PRGR_MOUSE     STANDARD;      PRT;    923 AA.
AC   Q00175;
DT   01-DEC-1992 (REL. 24, CREATED)
DT   01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   PROGESTERONE RECEPTOR (PR).
GN   PGR OR PR.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 91299759.
RA   SCHOTT D.R., SHYAMALA G., SCHNEIDER W., PARRY G.;
RL   BIOCHEMISTRY 30:7014-7020(1991).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; M68915; G200472; -.
DR   PIR; A39596; A39596.
DR   HSSP; P06536; 1GDC.
DR   PROSITE; PS00031; STEROID_FINGER.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING, NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
FT   DOMAIN         1     556       MODULATING, PRO-RICH.
FT   DNA_BIND     557     622       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING      557     577       C4-TYPE.
FT   ZN_FING      593     617       C4-TYPE.
FT   DOMAIN       671     923       STEROID-BINDING.
FT   DOMAIN       184     188       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL)
SQ   SEQUENCE   923 AA,  99073 MW,  EEDFF41A CRC32,

  Query Match             21.9%;  Score 681;  DB 7;  Length 923;
  Best Local Similarity  30.8%;  Pred. No. 3.86e-129;
  Matches  123;  Conservative  112;  Mismatches 138;  Indels  27;  Gaps  27;

Db     491 aapgtaaapaiyqplg-lnglpqlg-y-qaavlkdslpqvyppylnylrpdseasqspqy 547
            ::||  ::   . . |. | : ||  .    .|  |   .||  |     |.  :        ||
Qy      28 PTPGHLSPLVVHPQLSHLYAEPQKSPWCEAPSLEHTLP-VNPETLK-RKVSGNRCASPVT 85

Db     548 gfdslpq-kiclicgdeasgchygvltcgsckvffkramegqhnylcagrndcivdkirr 606
            :  |   :   :|  :|:| ||| |||| :| :|| ||||...|...|.|.. |.| .|| ||
        86 GPGSKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRR 145

Db     607 kncpacrlrkccqagmvl-ggrkfk-kfnkvrvmrtldgvalpqsvglpnesqalsqrit 664
            :|| ||||||| : ||| |:|:  .  .  ||  |. |   |    |  .  .   .
Qy     146 KSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQRSADE-QL-HCAGKAKRSGGHAPRVR 203

Db     665 fspnqeiqlvp-plinllmsiepdviyaghdntkpdtssslltslnqlgerqllsvvkws 723
                 :  |  |  |:  |:  ||  :  ::  . | | .|.. ||  |.....|. .. |.
Qy     204 -ELLLDA-LSPEQLVLTLLEAEPPHVLISRPSA-PFTEASMMMSLTKLADKELVHMISWA 260

Db     724 kslpgfrnlhiddqitliqyswmslmvfglgwrsykhvsgqmlyfapdliln-eq-r-mk 780
            | :||| :| : ||: |::  || ::: || |||   | :| :: |||||:|: :: : :
Qy     261 KKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVE 318

Db     781 e-lsfyslcltmwqipqefvklqvtheeflcmkvllllnt-i-pleglrsqsqfeemrss 837
             | :: : | :        |  | : | |:|||:| ::|||: : ||       :::   : :
Qy     319 GILEIFDM-L-L-ATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAH 375

Db     838 yirelikaig-lrqkgvvpssqrfyqltklldslhdlvkq 876
              :   :  |:  :  |: ::: |:   :|::||   |    |::
Qy     376 LLNAVTDALVWVIAKSGISSQQQSMRLANLL-MLLSHVRH 414


RESULT   12
ID   PRGR_HUMAN     STANDARD;       PRT;   933 AA.
AC   P06401;
DT   01-JAN-1988 (REL. 06, CREATED)
DT   01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   PROGESTERONE RECEPTOR (PR).
GN   PGR.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 90228361.
RA   KASTNER P., KRUST A., TURCOTTE B., STROPP U., TORA L., GRONEMEYER H.,
RA   CHAMBON P.;
RL   EMBO J. 9:1603-1614(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 87184565.
RA   MISRAHI M., ATGER M., D'AURIOL L., LOOSFELT H., MERIEL C.,
RA   FRIDLANSKY F., GUIOCHON-MANTEL A., GALIBERT F., MILGROM E.;
RL   BIOCHEM. BIOPHYS. RES. COMMUN. 143:740-748(1987).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- ALTERNATIVE PRODUCTS: THERE ARE TWO FORMS OF THE PROGESTERONE
CC       RECEPTOR: FORM A (THAT STARTS WITH MET-1) AND FORM B (THAT STARTS
```

```
CC       WITH MET-165).
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; X51730; G35652; -.
DR   EMBL; M15716; G189935; -.
DR   PIR; A03245; QRHUP.
DR   PIR; S09971; S09971.
DR   HSSP; P06536; 1GDC.
DR   MIM; 264080; -.
DR   PROSITE; PS00031; STEROID_FINGER.
DR   TRANSFAC; T00696; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING; PHOSPHORYLATION; ALTERNATIVE SPLICING.
FT   DOMAIN         1     566       MODULATING, PRO-RICH.
FT   DNA_BIND     567     632       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING      567     587       C4-TYPE.
FT   ZN_FING      603     627       C4-TYPE.
FT   DOMAIN       681     933       STEROID-BINDING.
FT   DOMAIN       183     187       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT   MOD_RES       41      41       PHOSPHORYLATION (BY CK).
FT   MOD_RES      227     227       PHOSPHORYLATION (BY CK).
FT   MOD_RES      232     232       PHOSPHORYLATION (BY CK).
FT   MOD_RES      552     552       PHOSPHORYLATION (BY CK).
FT   MOD_RES      793     793       PHOSPHORYLATION (BY CK).
FT   VARSPLIC       1     164       MISSING (IN FORM B).
FT   CONFLICT     226     226       S -> G (IN REF. 2).
FT   CONFLICT     256     256       C -> V (IN REF. 2).
FT   CONFLICT     344     344       T -> S (IN REF. 2).
FT   CONFLICT     660     660       L -> V (IN REF. 2).
SQ   SEQUENCE   933 AA,  99042 MW,  CD1124EB CRC32;

  Query Match             21.4%;  Score 666;  DB 7;  Length 933;
  Best Local Similarity  33.4%;  Pred. No. 1.67e-125;
  Matches  110;  Conservative   89;  Mismatches 113;  Indels  17;  Gaps  17;

Db     566 icliogdeasgchygvltcgsckvffkramegqhnylcagrndcivdkirrkncpacrlr 625
            :| :|:| ||| |||| |   || ||||    :|   |:|:: |:| :|| |||:| |||||
Qy      95 FCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRLR 154

Db     626 kccqagmvl-ggrkfk-kfnkvrvvraldavalpqplgvpnesqalsqrftfspgqdiql 683
            || : |||  |:|: :  :  ||  |: |    | :  | :: | : : |        |  |
Qy     155 KCYEVGMVKCGSRRERCGYRLVRRQRSADE-QL-HCAGKAKRSGGHAPRVR-ELLLDA-L 210

Db     684 ip-plinllmsiepdviyaghdntkpdtssslltslnqlgerqllsvvkwskslpgfrnl 742
              |  |:  |:  ||  :  ::  :: | | :|:: ||  |::::|: :: |:| :||| :|
Qy     211 SPEQLVLTLLEAEPPHVLISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVEL 269

Db     743 hiddqitliqyswmslmvfglgwrsykhvsgqmlyfapdlilneqrmk-essfyslcltm 801
            . ||. |..  || ..: || ||| | :| :: |||||:|: :  |   ::  :   :
Qy     270 SLFDQVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDML 327

Db     802 wqipqefvklqvsqeeflcmkvllllntiplegqlrsqtq-feemrs-syirelikaig- 857
                  |  | : : |:||:| ::|||: ::  | : ||  :  |  : :   :  |:
Qy     328 LATTSRFRELKLQHKEYLCVKAMILLNS-SMYPLVTATQDADSSRKLAHLLNAVTDALVW 386

Db     858 lrqkgvvsssqrfyqltklldnlhdlvkq 886
            :  |. :|| |:  .|..||  |   |::
Qy     387 VIAKSGISSQQQSMRLANLL-MLLSHVRH 414


RESULT   13
ID   PRGR_RABIT     STANDARD;       PRT;   930 AA.
AC   P06186;
DT   01-JAN-1988 (REL. 06, CREATED)
DT   01-JAN-1988 (REL. 06, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   PROGESTERONE RECEPTOR (PR).
GN   PGR.
OS   ORYCTOLAGUS CUNICULUS (RABBIT).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; LAGOMORPHA.
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RX   MEDLINE; 87067449.
RA   LOOSFELT H., ATGER M., MISRAHI M., GUIOCHON-MANTEL A., MERIEL C.,
RA   LOGEAT F., BENAROUS R., MILGROM E.;
RL   PROC. NATL. ACAD. SCI. U.S.A. 83:9045-9049(1986).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; M14547; G165632; -.
DR   PIR; A25923; A25922.
DR   HSSP; P06536; 1GDC.
DR   PROSITE; PS00031; STEROID_FINGER.
DR   TRANSFAC; T00697; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
FT   DOMAIN        1    565       MODULATING, PRO-RICH.
FT   DNA_BIND    568    633       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     568    588       C4-TYPE.
FT   ZN_FING     604    628       C4-TYPE.
FT   DOMAIN      678    930       STEROID-BINDING.
FT   DOMAIN      184    188       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
SQ   SEQUENCE   930 AA;  98666 MW;  FE133E86 CRC32;

  Query Match             21.3%;  Score 662;  DB 7;  Length 930;
  Best Local Similarity   32.5%;  Pred. No. 1.55e-124;
  Matches  106;  Conservative   91;  Mismatches  114;  Indels   15;  Gaps   15;

Db      567 kclicgdeasgchygvltcgsckvffkramegqhnylcagrndcivdkirrkncpacrlr 626
             :| :|:| ||| |||| :| :|| ||||:::|:::|:|:: |:| :|| |||:| |||||
Qy       95 FCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCIIDKNRRKSCQACRLR 154

Db      627 kccqaqmvl-ggrkfk-kfnkvrvmraldavalpqpvgipnesq-rltfspsqeiq-llp 682
            || : |||  |:|: :   :  ||  |: |   | :  | :· |  ·         ·: | |
Qy      155 KCYEVGMVKGGSRRERCGYRIVRRQRSADE-QL-HCAGKAKRSGGHAPRVRELLLDALSP 212

Db      683 -plinlimsiepdviyaqhdntkpdtsssiltslnqlgerqllsvvkwskslpgfrnlhi 741
              |:  |:  ||  .  .. ..  | | .|.. ||  |....|.    | | ·||| ·| ·
Qy      213 EQLVLTLLEAEPPHVLISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSL 271

Db      742 ddqitliqyswmslmvfglqwrsykhvsgqmlyfapdlilneqrmk-essfyslcltmwq 800
              ||: |::  || ::: || |||  | :| :: |||||:|: :  |   ::  :   :
Qy      272 FDQVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDMLLA 329

Db      801 ipqefvklqvsqeeflcmkvllllnt-i-pleqlrsqsqfeemrssyirelikaig-lrq 857
                |  | : : |:||:| ::|||: : ||      :::    : :  :  : |: :
Qy      330 TTSRFRELKLQHKEYLCVKAMILLNSSMYPLVIAIQDADSSRKLAHLLNAVIDALVWVIA 389

Db      858 kqvvsssqrfyqltklldnlhdlvkq 883
            |: :|| |:  :|::||  |   |::
Qy      390 KSGISSQQQSMRLANLL-MLLSHVRH 414


RESULT     14
ID   GCR_HUMAN      STANDARD;      PRT;   777 AA.
AC   P04150; P04151;
DT   01-NOV-1986 (REL. 03, CREATED)
DT   01-NOV-1986 (REL. 03, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   GLUCOCORTICOID RECEPTOR (GR).
GN   GRL.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 86092296.
RA   HOLLENBERG S.M., WEINBERGER C., ONG E.S., CERELLI G., ORO A.,
RA   LEBO R., THOMPSON E.B., ROSENFELD M.G., EVANS R.M.;
RL   NATURE 318:635-641(1985).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   MUNROE D.G., PANG J., TAYLOR G.R., LAU C., PLANTE R.K., ZHOU L.;
RL   SUBMITTED (SEP-1993) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [3]
RP   SEQUENCE OF 1-394 FROM N.A.
RX   MEDLINE; 91224961.
RA   LECLERC S., XIE B.X., ROY R., GOVINDAN M.V.;
RL   J. BIOL. CHEM. 266:8711-8719(1991).
RN   [4]
RP   DOMAINS.
RX   MEDLINE; 86092211.
RA   WEINBERGER C., HOLLENBERG S.M., ROSENFELD M.G., EVANS R.M.;
RL   NATURE 318:670-672(1985).
RN   [5]
RP   VARIANT SER-363.
RX   MEDLINE; 93187003.
RA   KARL M., LAMBERTS S.W.J., DETERA-WADLEIGH S.D., ENCIO I.J.,
RA   STRATAKIS C.A., HURLEY D.M., ACCILI D., CHROUSOS G.P.;
RL   J. CLIN. ENDOCRINOL. METAB. 76:683-689(1993).
RN   [6]
RP   VARIANT VAL-641.
RX   MEDLINE; 91123468.
RA   HURLEY D.M., ACCILI D., STRATAKIS C.A., KARL M., VAMVAKOPOULOS N.,
RA   RORER E., CONSTANTINE K., TAYLOR S.I., CHROUSOS G.P.;
RL   J. CLIN. INVEST. 87:680-686(1991).
RN   [7]
RP   VARIANT ILE-729.
RX   MEDLINE; 93253031.
RA   MALCHOFF D.M., BRUFSKY A., REARDON G., MCDERMOTT P., JAVIER E.C.,
RA   BERGH C.H., ROWE D., MALCHOFF C.D.;
RL   J. CLIN. INVEST. 91:1918-1925(1993).
RN   [8]
RP   VARIANT PHE-753.
RX   MEDLINE; 93302771.
RA   ASHRAF J., THOMPSON E.B.;
RL   MOL. ENDOCRINOL. 7:631-642(1993).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES. BINDS TO THE
CC       GRE TARGET SITE.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
CC   -!- ALTERNATIVE PRODUCTS: ALPHA AND BETA FORMS OF HGR DIFFER ONLY IN
CC       THEIR CARBOXYL-TERMINAL SEQUENCES. THEY ARE GENERATED BY
CC       ALTERNATIVE SPLICING. ALPHA-HGR IS THE PREDOMINANT PHYSIOLOGICAL
CC       FORM OF HGR. THE IN VITRO TRANSLATION PRODUCT OF HGR BETA DOES NOT
CC       BIND STEROID. ITS FUNCTION IS UNCLEAR; HOWEVER, IT IS POSSIBLE THAT
CC       VARIANT RECEPTORS PERFORM TISSUE-SPECIFIC FUNCTIONS.
CC   -!- DISEASE: DEFECTS IN GRL ARE A CAUSE OF FAMILIAL GLUCOCORTICOID
CC       RESISTANCE (CORTISOL RESISTANCE): A HYPERTENSIVE, HYPERANDROGENIC
CC       DISORDER CHARACTERIZED BY INCREASED SERUM CORTISOL CONCENTRATIONS.
DR   EMBL; M10901; G183033; -.
DR   EMBL; M11050; G183035; -.
DR   EMBL; U01351; G458657; -.
DR   EMBL; X03225; G31680; -.
DR   EMBL; X03348; G31682; -.
DR   EMBL; M69104; G553322; -.
DR   EMBL; M73816; G553323; -.
DR   EMBL; S68378; G239758; -.
DR   PIR; A03246; QRHUGA.
DR   PIR; A03247; QRHUGB.
DR   HSSP; P06536; 1GDC.
DR   MIM; 138040; -.
DR   PROSITE; PS00031; STEROID_FINGER.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING; DISEASE MUTATION; POLYMORPHISM;
KW   ALTERNATIVE SPLICING.
FT   DOMAIN      145    278       IMMUNOGENIC.
FT   DNA_BIND    421    486       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     421    441       C4-TYPE.
FT   ZN_FING     457    491       C4-TYPE.
FT   DOMAIN      487    527       HINGE.
FT   DOMAIN      528    777       STEROID-BINDING.
FT   VARIANT     363    363       N -> S.
FT   VARIANT     641    641       D -> V (IN GLUCOCORTICOID RESISTANCE).
FT   VARIANT     729    729       V -> I (IN GLUCOCORTICOID RESISTANCE).
FT   VARIANT     753    753       L -> F (IN GLUCOCORTICOID RESISTANCE).
FT   VARSPLIC    728    777       VVENLLNYCFQTFLDKTMSIEFPEMLAEIITNQIPKYSNGN
FT                                IKKLLFHQK -> ENVMWIKPESTSHTLI (IN BETA
FT                                FORM).
FT   CONFLICT    451    451       G -> GR (IN REF. 2)
SQ   SEQUENCE   777 AA;  85659 MW;  D93D1F01 CRC32;

  Query Match             19.8%;  Score 614;  DB 4;  Length 777;
  Best Local Similarity  33.4%;  Pred. No. 5.96e-113;
  Matches    97;  Conservative   84;  Mismatches  93;  Indels  16;  Gaps  15;

Db     397 smrpdvsspppssstattgp-p-pkl-lvcsdeasgchygvltcgsckvffkraveqqhn 454
           :::  ||:   :|: :  |:    : :| |||| ||| |||| :| :|| ||||::|:::
        70 TLKRKVSGNRCASPVTGPGSKRDAHRCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHND 129

Db     455 ylcagrndcildkirrkncpacryrkclqagmnl-earktk-k-kikgiqqattg-v-sq 509
           |:|:: |:| ||| |||:| ||| ||| : ||     :|: .   ::   |...   :
Qy     130 YICPATNQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCA 189

Db     510 etsenpgnkt-ivpatlpq-ltpt-lv-sllviepevlyagydssvpdstwrimttlnm 565
            .  :|  .  '   ' . |:|  |' :'|'   ' |'  .    .   :   |  '
Qy     190 GKAKRSGGHAPRVRELLLDALSPEQLVLTLLEAEPPHVLISRPSAPFTEAS--MMMSLTK 247

Db     566 lggrqviaavkwakaipgfrnlhlddqmtllqyswmflmafalgwrsyrqssanllcfap 625
           |: ::::  : ||| |||| :| | ||  ||   ||       .| |||   .:. :|: |||
Qy     248 LADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPG-KLI-FAP 305

Db     626 dliin-eqrmtlpcmydqckhmlyvsselhrlqvsyeeylcmktlllls 674
           ||::: ::  .  .   .   .:   .|    |      ||||:|    ||:|
Qy     306 DLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNS 355


RESULT   15
ID   GCR_XENLA      STANDARD;      PRT;   776 AA.
AC   P49844;
DT   01-OCT-1996 (REL. 34, CREATED)
DT   01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   GLUCOCORTICOID RECEPTOR (GR).
OS   XENOPUS LAEVIS (AFRICAN CLAWED FROG).
     EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AMPHIBIA; ANURA.
     [1]
     SEQUENCE FROM N.A.
RC   TISSUE=EMBRYO;
RX   MEDLINE; 94289478.
RA   GAO X., KALKHOVEN E., PETERSON-MADURO J., VAN DER BURG B.,
RA   DESTREE O.H.J.;
RL   BIOCHIM. BIOPHYS. ACTA 1218:194-198(1994).
RN   [2]
RP   SEQUENCE OF 452-776 FROM N.A.
RC   TISSUE=LIVER;
RA   PICARD D.;
RL   SUBMITTED (FEB-1994) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES. BINDS TO THE
CC       GRE TARGET SITE.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; X72211; G444043; -.
DR   EMBL; X77764; G456374; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
```

```
FT   DOMAIN        1    419       MODULATING.
FT   DNA_BIND    420    485       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     420    440       C4-TYPE.
FT   ZN_FING     456    480       C4-TYPE.
FT   DOMAIN      486    776       STEROID-BINDING.
FT   CONFLICT    502    502       S -> P (IN REF. 2).
FT   CONFLICT    504    504       T -> A (IN REF. 2).
FT   CONFLICT    506    506       T -> A (IN REF. 2).
FT   CONFLICT    511    511       P -> N (IN REF. 2).
FT   CONFLICT    519    519       L -> M (IN REF. 2).
FT   CONFLICT    544    544       Y -> F (IN REF. 2).
FT   CONFLICT    551    551       I -> M (IN REF. 2).
FT   CONFLICT    654    654       S -> R (IN REF. 2).
FT   CONFLICT    740    740       L -> M (IN REF. 2).
FT   CONFLICT    768    768       L -> I (IN REF. 2).
SQ   SEQUENCE   776 AA;  84977 MW;  8960496D CRC32;

  Query Match             19.6%;  Score 608;  DB 4;  Length 776;
  Best Local Similarity  34.9%;  Pred. No. 1.65e-111;
  Matches    99;  Conservative   74;  Mismatches  96;  Indels  15;  Gaps  15;

Db     398 irsdaspspstsststgpppklclvcsdeasgchygvltcgsckvffkraveqqhnylca 457
           : :.  :|| |.. |      : .| |||| ||| |||| .| .|| |||.::|:::|:|:
Qy      75 VSGNRCASPVTGPGSK-RDAHRCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICP 133

Db     458 grndciidkirrkncpacryrkclqagmnl-earktk-k-kikgiqqsttatar-es-pe 512
           : |.| ||| |||:| ||| ||| : ||     :|: :    ::    |:|:     : : :
Qy     134 ATNQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAK 193

Db     513 tsmtrtl-vpasv-aqltpt-li-sllviepevlysgydssipdttrrlmsslnmlqgr 568
           |  :.  |   :   |.|  |: .|||   | || |   ... ...   .|| | :|: :
Qy     194 RSGGHAPRVRELLLDALSPEQLVLTLLEAEPPHVLISRPSAPFTEASM-MMS-LTKLADK 251

Db     569 qvvsavrwakaipgfrnlhlddqmtllqyswmflmvfalgwrsykqtngsllyfapdlvl 628
           .:|  . ||| |||| :| | ||. ||. || ... .| |||  :   '   | |||||||:
Qy     252 ELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHP-GKLIFAPDLVL 309

Db     629 t-edrmhlpfmqercqemlkiagemsslqisydeylcmkvlllm 671
            .. . .  . | : .|  .:    | :   ||||:| ::|:
Qy     310 DRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILL 353


Search completed. Thu Dec 18 08.57.11 1997
Job time : 43 secs.
```

```
**************************************************************************
  ___    ___    ____    ____    ____    ____    _    _
 |   \  /   |  |  _ \  / ___|  |  _ \  / ___|  | |  | |
 | |\ \/ /| |  | |_) | \___ \  | |_) || |      | |__| |
 | | \__/ | |  |  __/   ___) | |  _ < | |      |  __  |
 | |      | |  | |     |     | | | \ \| |___   | |  | |
 |_|      |_|  |_|     |____/  |_|  \_\\____|  |_|  |_| (TM)

**************************************************************************

          Release 2 1D John F Collins, Biocomputing Research Unit
          Copyright (c) 1993, 1994, 1995 University of Edinburgh, U K
                 Distribution rights by IntelliGenetics, Inc

MPsrch_pp   protein - protein database search, using Smith-Waterman algorithm

  on:             Thu Dec 18 09:54:13 1997   MasPar time 15 22 Seconds
                                             789 706 Million cell updates/sec
abular output not generated

Title:            >US-08-826-361A-6
Description:      (1-416) from US08826361A.pep
Perfect Score:    3108
Sequence:         1 MNYSIPSNVINLEGGPGRQT          QQSMRIANLLMLLSHVRHAR 416

Scoring table:    PAM 150
                  Gap 11

Searched:         91006 seqs, 28888923 residues

Post-processing:  Minimum Match 0%
                  Listing first 45 summaries

Database:         pir51
                    1:ann1 2:ann2 3:ann3 4:ann4 5:unann1 6:unann2 7:unann3
                    8:unann4 9:unann5 10:unann6 11:unann7 12:unann8
                    13:unann9 14:unann10 15:unenc 16:unrev

Statistics:       Mean 47.477;  Variance 101.166;  scale 0.469

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                  SUMMARIES
                %
esult           Query
  No.   Score   Match Length DB  ID              Description              Pred. No.
-------------------------------------------------------------------------------------
    1    1512    48.6    586  3  QRXLE           estrogen receptor -     2.48e-263
    2    1484    47.7    589  3  QRCHE           estrogen receptor -     8.49e-258
    3    1470    47.3    701 16  S64737          80K estrogen recepto    4.95e-255
    4    1463    47.1    595  3  QRHUE           estrogen receptor -     1.20e-253
    5    1455    46.8    599  3  QRMSE           estrogen receptor -     4.55e-252
    6    1449    46.6    600  3  QRRTE           estrogen receptor -     1.10e-250
    7    1327    42.7    574  7  A37197          estrogen receptor -     7.87e-227
    8    1314    42.3    595  7  S58224          oestrogen receptor -    2.85e-224
    9     740    23.8    473 14  S58097          estrogen receptor re    1.41e-112
   10     736    23.7    433 13  B29345          steroid hormone rece    9.21e-112
   11     725    23.3    521 13  A29345          steroid hormone rece    1.04e-109
   12     681    21.9    923  7  A39946          progesterone recepto    2.56e-101
   13     671    21.6    923 14  I53280          progesterone recepto    2.04e-99
   14     666    21.4    933  3  QRHUP           progesterone recepto    1.82e-98
   15     662    21.3    930  7  A25923          progesterone recepto    1.05e-97
   16     614    19.8    742  3  QRHUGB          glucocorticoid recep    1.29e-88
   17     614    19.8    777  3  QRHUGA          glucocorticoid recep    1.29e-88
   18     608    19.6    776  3  S44047          glucocorticoid recep    1.75e-87
   19     597    19.2    786  7  A35466          progesterone recepto    2.07e-85
   20     594    19.1    771  7  A54273          glucocorticoid recep    7.61e-85
   21     586    18.9    783  3  A25691          glucocorticoid recep    2.44e-83
   22     579    18.6    133  7  S26595          estrogen receptor -     5.05e-82
   23     577    18.6    795  3  QRRTG           glucocorticoid recep    1.20e-81
   24     562    18.1    981 14  A41401          mineralocorticoid re    7.83e-79
   25     557    17.9    899 14  A45895          androgen receptor -     6.77e-78
   26     554    17.8    902 14  B47494          androgen receptor -     1.04e-77
   27     553    17.8    919 13  A39249          androgen receptor -     3.80e-77
   28     550    17.7    910 13  A34721          androgen receptor A     1.38e-76
   29     550    17.7    911 13  B34721          androgen receptor B     1.38e-76
   30     544    17.5    759 16  S60536          glucocorticoid recep    1.83e-75
   31     541    17.4    344 13  I51331          androgen receptor -     1.12e-74
   32     523    16.8    984 13  A29513          mineralocorticoid re    1.52e-71
   33     506    16.3    533 13  S37781          retinoid X receptor     2.20e-68
   34     505    16.2    410 14  S26669          retinoic acid recept    3.37e-68
   35     505    16.2    446 14  A44419          H-2 region II bindin    3.37e-68
   36     505    16.2    448 14  D41727          retinoid X receptor     3.37e-68
   37     505    16.2    451 14  A41651          retinoic acid recept    3.37e-68
   38     501    16.1    563  7  JQ0623          nerve growth factor-    1.86e-67
   39     493    15.9    462 13  S03532          retinoic acid recept    5.64e-66
   40     494    15.9    466 14  C41727          retinoid X receptor     3.69e-66
   41     494    15.9    467 14  S16669          retinoic acid recept    3.69e-66
   42     494    15.9    467 14  A47278          retinoid X receptor     3.69e-66
   43     494    15.9    598  7  A37251          probable nuclear hor    2.41e-66
   44     490    15.8    598  7  A61451          steroid hormone rece    2.03e-65
   45     489    15.7    601  3  QRMSN1          probable hormone rec    3.10e-65

                                 ALIGNMENTS

RESULT      1
ENTRY                QRXLE      #type complete
TITLE                estrogen receptor - African clawed frog
ORGANISM             #formal_name Xenopus laevis #common_name African clawed frog
DATE                 28-Feb-1992 #sequence_revision 14-Jul-1994 #text_change
                       15-Jun-1996
ACCESSIONS           A40907
REFERENCE            A40907
   #authors          Weiler, I.J.; Lew, D.; Shapiro, D.J.
   #journal          Mol. Endocrinol. (1987) 1:355-362
   #title            The Xenopus laevis estrogen receptor: sequence homology with
                       human and avian receptors and identification of multiple
                       estrogen receptor messenger ribonucleic acids.
   #cross-references MUID:90331927
   #accession        A40907
      ##molecule_type mRNA
      ##residues       1-586 ##label WEI
      ##cross-references GB:L20735
COMMENT     The steroid hormones and their receptors are involved in the
              regulation of eukaryotic gene expression and affect cellular
              proliferation and differentiation in target tissues.
COMMENT     In the absence of ligand, steroid hormone receptors are thought to
              be weakly associated with nuclear components; hormone binding
              greatly increases receptor affinity. The hormone-receptor complex
              appears to recognize discrete DNA sequences upstream of
              transcriptional start sites.
CLASSIFICATION   #superfamily estrogen receptor, erbA transforming protein
                   homology
KEYWORDS         DNA binding; nucleus; phosphoprotein; steroid binding;
                   steroid hormone receptor; transcription regulation; zinc
                   finger
FEATURE
   1-174               #domain amino-terminal #label NH2\
   175-266             #domain DNA binding #status predicted #label DNA\
   178-448             #domain erbA transforming protein homology #label ERBA\
   178-201             #region zinc finger C4C4 motif\
   214-237             #region zinc finger C2C2 motif\
   251-266             #region nuclear location signal\
   292-544             #domain steroid binding #status predicted #label STE\
   180,183,197,200     #binding_site zinc (Cys) #status predicted\
   216,222,232,235     #binding_site zinc (Cys) #status predicted\
   231                 #binding_site phosphate (Ser) (covalent) #status
                         predicted
SUMMARY          #length 586  #molecular-weight 66080  #checksum 9991
```

```
  Query Match             48.6%;  Score 1512;  DB 3;  Length 586;
  Best Local Similarity   52.1%;  Pred. No. 2.48e-263;
  Matches  213;  Conservative   95;  Mismatches  84;  Indels  17;  Gaps  13;

Db     101 spvvflaklpqlspflhhhqqqvpyylesqqtf-avreaapptfyrsssdnrrqsgrer 159
           || |:  :   :||||:: |: |      | |:: :        |         .  :: || ::
Qy      22 SPNVLWPTPGHLSPLVVHR-QLSHLYAEPQKSPWCEARSLEHTLPVNRETLKRKVSG-NR 79

Db     160 mssandkqppsmestketrycavcsdyasgyhygvwscegckaffkrsiqghndymcpat 219
              ::    || :   | :::::|||||||||||||||||||||||||||||||||.||||
Qy      80 CASPVTGP-G--SKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAT 135

Db     220 nqctidknrrkscqacrlrkcyevgmmkqqirkdrrqqrllkhkrqkeeqeqkndvdpse 279
           |||||||||||||||||||||||||||||:| | |:.| | ||..  ||.  .||  .
Qy     136 NQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVR--RQRSADEQLH-CAGKA 192

Db     280 irtasiwvnpsvksmklspvlsltaeqlisalmeaeapivysehdstkplseasmmtllt 339
            |:::     | |: : | : ||   :|||: :|:||||:| |     : |: |::||||  ||
Qy     193 KRSGG--HAPRVRELLLDA-LS--PEQLVLTLLEAEPPHVLISRPSA-PFTEASMMMSLT 246

Db     340 nladrelvhminwakrvpgfvdltlhdqvhllecawleilmvgliwrsvehpgklsfapn 399
           :|||:|||||||:||||::||||:|:| |||:|||   |:|:||:||:||::|||||  |||:
Qy     247 KLADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPD 306

Db     400 llldrnqgrcveglveifdmlvttatrfrmmrlrqeeficlksiillnsqvytflsstle 459
           |:|||::|:||||::|||||||::|::|||  ::|.   |..|.|.  ::::: |  :    |
Qy     307 LVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQD 366

Db     460 sledtdlihiildkiidtlvhfmaksqlslqqqqrrlaqllllilshirh 508
           :       | |: |: : |:||  :||||:| |||  ||| ||::|||:||
Qy     367 ADSSRKLAHL-LNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRH 414


RESULT     2
ENTRY                 ORCHE      #type complete
TITLE                 estrogen receptor - chicken
ORGANISM              #formal_name Gallus gallus #common_name chicken
DATE                  06-Mar-1992 #sequence_revision 14-Jul-1994 #text_change
                        15-Jun-1996
ACCESSIONS            A40914; S07192
REFERENCE             A40914
   #authors           Maxwell, B.L.; McDonnell, D.P.; Conneely, O.M.; Schulz, T.Z.;
                        Greene, G.L.; O'Malley, B.W.
   #journal           Mol. Endocrinol. (1987) 1:25-35
   #title             Structural organization and regulation of the chicken
                        estrogen receptor.
   #cross-references MUID:88318621
   #accession         A40914
      ##status              not compared with conceptual translation
      ##molecule_type mRNA
      ##residues            1-589 ##label MAX
REFERENCE             S07192
   #authors           Krust, A.; Green, S.; Argos, P.; Kumar, V.; Walter, P.;
                        Bornert, J.M.; Chambon, P
   #journal           EMBO J. (1986) 5:891-897
   #title             The chicken oestrogen receptor sequence: homology with v-erbA
                        and the human oestrogen and glucocorticoid receptors.
   #cross-references MUID:86247578
   #accession         S07192
      ##status              preliminary
      ##molecule_type mRNA
      ##residues            1-256,'R',258-589 ##label KRU
      ##cross-references EMBL:X03805
COMMENT    The steroid hormones and their receptors are involved in the
             regulation of eukaryotic gene expression and affect cellular
             proliferation and differentiation in target tissues.
COMMENT    In the absence of ligand, steroid hormone receptors are thought to
             be weakly associated with nuclear components; hormone binding
             greatly increases receptor affinity. The hormone-receptor complex
             appears to recognize discrete DNA sequences upstream of
             transcriptional start sites
CLASSIFICATION   #superfamily estrogen receptor; erbA transforming protein
                        homology
KEYWORDS              DNA binding; nucleus; phosphoprotein; steroid binding;
                        steroid hormone receptor; transcription regulation; zinc
                        finger
FEATURE
   1-173                    #domain amino-terminal #label NH2\
   174-265                  #domain DNA binding #status predicted #label DNA\
   177-450                  #domain erbA transforming protein homology #label ERBA\
   177-200                  #region zinc finger CCCC motif\
   213-235                  #region zinc finger CCCC motif\
   250-265                  #region nuclear location signal\
   294-546                  #domain steroid binding #status predicted #label STR\
   179,182,196,199          #binding_site zinc (Cys) #status predicted\
   215,221,231,234          #binding_site zinc (Cys) #status predicted\
   230,299                  #binding_site phosphate (Ser) (covalent) #status
                              predicted
SUMMARY               #length 589  #molecular-weight 66773  #checksum 793

  Query Match             47.7%;  Score 1484;  DB 3;  Length 589;
  Best Local Similarity   52.1%;  Pred. No. 8.49e-258;
  Matches  214;  Conservative   95;  Mismatches  84;  Indels  18;  Gaps  11;

Db     100 spvvflqtapqlspflhhhsqqvpyylenepsgfmreaappafyrpssdnrrhsirerm 159
           || |:  |: :|||:: |: |      | | : ::      :    ::     : :   :: :
Qy      22 SPNVLWPTPGHLSPLVVHR-QLSHLYAEPQKSPWCEARSLEHTL-PVNRETLKRKVSGNR 79

Db     160 sstndkaslsmestketrycavcndyasgyhygvwscegckaffkrsiqghndymcpatn 219
            ::    |: :   | :::::||||:|||||||||||||||||||||||||||||||:|||||
Qy      80 CASPVTGP-G--SKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATN 136

Db     220 qctidknrrksqqacrlrkcyevgmmkggirkdrrggrmlkhkrqrdeqdsrneaasste 279
           ||||||||||||||||||||||||||||:| | |::| | |:::: |  :||     | |  :
Qy     137 QCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSG 196

Db     280 lraptlwtsplvvkhnkknspalsltaeqmvsalleaeppivysoydpnrpfneasmmtl 339
            :|| ::   |::       :  |||  :||:| :|||||||| |       |: || |||||
Qy     197 GHAPRVR-E-LLL----D--ALS--PEQLVLTLLEAEPPHVLIS-RPSAPFTEASMMMS 244

Db     340 ltnladrelvhminwakrvpgfvdltlhdqvhllecawleilmiglvwrsmehpgkllfa 399
           ||:|||:|||||||:|||::||||:|:| |||:|||  |:|:||:||:|||::||||||:||
Qy     245 LIKLADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFA 304

Db     400 pnlllldrnqgkcvegmvelfdmllatsarfrmmnlqgeefvclksiillnsgvytflsst 459
           |:|:|||::|||||::||||||||||::|||  ::||   |::|:|::||||||::|  ::::|
Qy     305 PDLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTAT 364

Db     460 lksleerdyihrvldkitdtlihlmaksqlslqqqhrrlaqllilshirh 510
            :    |     | :|: :||:|: ::||||:| |||   ||| ||::||||:||
Qy     365 QDADSSRKLAH-LLNAVIDALVWVIAKSGISSQQQSMRLANLLMLLSHVRH 414


RESULT     3
ENTRY                 S64737      #type complete
TITLE                 80K estrogen receptor - human
ORGANISM              #formal_name Homo sapiens #common_name man
DATE                  28-Oct-1996 #sequence_revision 28-Oct-1996 #text_change
                        28-Oct-1996
ACCESSIONS            S64737
REFERENCE             S64737
   #authors           Pink, J.J.; Wu, S.Q.; Wolf, D.M.; Bilimoria, M.M.; Jordan,
                        V.C.
   #journal           Nucleic Acids Res. (1996) 24:962-969
   #title             A novel 80 kDa human estrogen receptor containing a
                        duplication of exons 6 and 7.
   #accession         S64737
      ##status              preliminary
      ##residues            1-701 ##label PIN
      ##cross-references EMBL:U47678
SUMMARY               #length 701  #molecular-weight 78430  #checksum 4951

  Query Match             47.3%;  Score 1470;  DB 16;  Length 701;
  Best Local Similarity   57.3%;  Pred. No. 4.95e-255;
  Matches  196;  Conservative   84;  Mismatches  47;  Indels  15;  Gaps  11;
```

```
Db     178 saketrycavcndyasgyhygvwscegckaffkrsiqghndymcpatnqctidknrrksc 237
           | :::::||||:|||||||||||||||||||||||||||||||:|||||||||||||||||||
Qy      89 SKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSC 148

Db     238 qacrlrkcyevgmmkggirkdrrggrmlkhkrqrddgegrgevgsagdmraanlwpsplm 297
           ||:|||:||||.|:| | |::| | |:::  |||    | :   :: : :|::    :|
Qy     149 QACRLRKCYEVGMVKCGSRRERCGYRLVR--RQRSADE-QLHCAGKA-KRSGG--HAP-R 201

Db     298 ikrskknslalsltadqmvsalldaepp-ilyseydptrpfseasmmglltnladrelvh 356
           : |     |  :|:::|:| :||:|||| :| |   |. ||.||||  ||:|||:||||
Qy     202 V-RELL--LD-AISPEQLVLTLLEAEPPHVLISR--PSAPFTEASMMMSLTKLADKELVH 255

Db     357 minwakrvpgfvdltlhdqvhllecawleilmiglvwrsmehpvkllfapnllldrnqgk 416
           ||:|||::||||:|:| |||:|||  |:|:||:||:|:|·:|| ||·|||·|·|||··||
Qy     256 MISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGK 315

Db     417 cvegmveifdmllatssrfrmmnlqgeefvclksiillnsgvytflsstlksleekdhih 476
           ||||::|||||||| |||| ..|| |:.|.|..|||||..| ...:|  : :    :
Qy     316 CVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSRKLA 374

Db     477 rvldkitdtlihlmakagltlqqqhqrlaqlllilshirhmr 518
           ··|· ·||·|· ··||·|   |||  ||| ||··|||·|| |
Qy     375 HLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHAR 416


RESULT     4
ENTRY               QRHUE       #type complete
TITLE               estrogen receptor - human
ALTERNATE_NAMES     hER
ORGANISM            #formal_name Homo sapiens #common_name man
DATE                28-May-1986 #sequence_revision 28-May-1986 #text_change
                      15-Nov-1996
ACCESSIONS          A94284; A93376; A43021; S27143; S34000; A41925; B41925;
                      A03244; C41925; D41925
REFERENCE           A94284
   #authors         Greene, G.L.; Gilna, P.; Waterfield, M.; Baker, A.; Hort, Y.;
                      Shine, J.
   #journal         Science (1986) 231:1150-1154
   #title           Sequence and expression of human estrogen receptor
                      complementary DNA
   #cross-references MUID:86122927
   #accession       A94284
      ##molecule_type mRNA
      ##residues          1-595 ##label GR1
      ##cross-references GB:M12674
REFERENCE           A93376
   #authors         Green, S.; Walter, P.; Kumar, V.; Krust, A.; Bornert, J.M.;
                      Argos, P.; Chambon, P.
   #journal         Nature (1986) 320:134-139
   #title           Human oestrogen receptor cDNA: sequence, expression and
                      homology to v-erb-A.
   #cross-references MUID:86146892
   #accession       A93376
      ##molecule_type mRNA
      ##residues          1-595 ##label GR2
      ##cross-references GB:X03635
REFERENCE           A43021
   #authors         Ponglikitmongkol, M.; Green, S.; Chambon, P.
   #journal         EMBO J. (1988) 7:3385-3388
   #title           Genomic organization of the human oestrogen receptor gene.
   #accession       A43021
      ##molecule_type DNA
      ##residues          143-181,205-225,244-264,356-374,[illegible] 412-422,442-461,
                            'P',462-467;509-528 ##label PON
REFERENCE           S27140
   #authors         Keaveney, M.; Klug, J.; Gannon, F.
   #journal         DNA Seq. (1992) 2:347-358
   #title           Sequence analysis of the 5' flanking region of the human
                      estrogen receptor gene.
   #accession       S27143
      ##status            preliminary; translation not shown
      ##molecule_type DNA
      ##residues          1-115 ##label KEA
```

```
      ##cross-references EMBL:X62462
REFERENCE           S34000
   #authors         Pfeffer, U.; Fecarotta, E.; Castagnetta, L.; Vidali, G.
   #journal         Cancer Res. (1993) 53:741-743
   #title           Estrogen receptor variant messenger RNA lacking exon 4 in
                      estrogen-responsive human breast cancer cell lines
   #accession       S34000
      ##status            preliminary
      ##molecule_type mRNA
      ##residues          216-254,367-399,'G',401-434 ##label PFE
      ##cross-references EMBL:X73067
      ##note              the authors translated the codon GGG for residue 400 as
                            Val
REFERENCE           A41925
   #authors         Dotzlaw, H.; Alkhalaf, M.; Murphy, L.C.
   #journal         Mol. Endocrinol. (1992) 6:773-785
   #title           Characterization of estrogen receptor variant mRNAs from
                      human breast cancers.
   #cross-references MUID:92293154
   #accession       A41925
      ##molecule_type mRNA
      ##residues          1-214,'ELPTLC' ##label DOT
      ##cross-references NCBIN:106580
      ##experimental_source clone 4; breast cancer
      ##note              sequence has been revised after extraction from NCBI
                            backbone
      ##note              the complete sequence of neither the nucleotide nor the
                            protein is shown in this paper
   #accession       B41925
      ##status            significant sequence differences
      ##molecule_type mRNA
      ##cross-references NCBIN:106597; GB:M69296
      ##experimental_source clone 24; breast cancer
COMMENT    The steroid hormones and their receptors are involved in the
             regulation of eukaryotic gene expression and affect cellular
             proliferation and differentiation in target tissues.
COMMENT    In the absence of ligand, steroid hormone receptors are thought to
             be weakly associated with nuclear components; hormone binding
             greatly increases receptor affinity. The hormone-receptor complex
             appears to recognize discrete DNA sequences upstream of
             transcriptional start sites.
GENETICS
   #gene            GDB:ESR
      ##cross-references GDB:119120
   #map_position 6q25.1-6q25.1
   #introns         151/2; 215/1; 254/1; 366/1; 412/2; 457/1; 518/2
CLASSIFICATION      #superfamily estrogen receptor; erbA transforming protein
                      homology
KEYWORDS            DNA binding; nucleus; phosphoprotein; steroid hormone
                      receptor; transcription regulation; zinc finger
FEATURE
   1-120                  #domain amino-terminal #label NH2\
   121-299                #domain DNA binding #status predicted #label DNA\
   183-456                #domain erbA transforming protein homology #label ERBA\
   185-205                #region zinc finger CCCC motif\
   221-245                #region zinc finger CCCC motif\
   300-595                #domain steroid binding #status predicted #label STR\
   236,305                #binding_site phosphate (Ser) (covalent) #status
                            predicted
SUMMARY             #length 595  #molecular-weight 66258  #checksum 4317

  Query Match             47.1%;  Score 1463;  DB 3;  Length 595;
  Best Local Similarity   57.4%;  Pred. No. 1.20e-253;
  Matches  135;  Conservative   84;  Mismatches  46;  Indels  15;  Gaps  11;

Db     178 saketrycavcndyasgyhygvwscegckaffkrsiqghndymcpatnqctidknrrksc 237
           | :::::||||:|||||||||||||||||||||||||||||||:|||||||||||||||||||
Qy      89 SKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSC 148

Db     238 qacrlrkcyevgmmkggirkdrrggrmlkhkrqrddgegrgevgsagdmraanlwpsplm 297
           |||||||||||||.| | |..| | |...  |||    | :   .. : :|::    :|
Qy     149 QACRLRKCYEVGMVKCGSRRERCGYRLVR--RQRSADE-QLHCAGKA-KRSGG--HAP-R 201
```

```
Db     298 lrskkn-lnlsltadqmvsalldaeppilyseydptrpfseasmmglltnladrelvh 356
           : |      |  :|:::|:| :||:|||| :| |    |: ||:|||||  ||:|||:||||
Qy     202 V-RFLL--LD-ALSPPQLVLTLLEAEPPHVLISR--PSAPFTEASMMMSLTKLADKELVH 255

Db     357 minwakrvpgfvdltlhdqvhllecawleilmiglvwrsmehpvkllfapnllldrnqgk 416
           ||:|||::||||:|:| |||:|||  |:|:||:||:|||::|| ||:|||:|:|||::||
Qy     256 MISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGK 315

Db     417 cveqmveifdmllatssrfrmmnlqgeefvclksiillnsgvytflsstlksleekdhih 476
           ||||::||||||||:|||| ::||  |::|:|::||||||::| ::::|  : :     :
Qy     316 CVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSRKLA 374

Db     477 rvldkitdtlihlmakagltlqqqhqrlaqllliishirh 516
           ::|: :||:|: ::||:|:: |||  ||| ||::|||:||
Qy     375 HLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRH 414
```

```
RESULT      5
ENTRY              QRMSE      #type complete
TITLE              estrogen receptor - mouse
ORGANISM           #formal_name Mus musculus #common_name house mouse
DATE               06-Mar-1992 #sequence_revision 14-Jul-1994 #text_change
                     15-Jun-1996
ACCESSIONS         A40061
REFERENCE          A40061
   #authors        White, R.; Lees, J.A.; Needham, M.; Ham, J.; Parker, M.
   #journal        Mol  Endocrinol  (1987) 1:735-744
   #title          Structural organization and expression of the mouse estrogen
                     receptor.
   #cross-references MUID:91042558
   #accession      A40061
      ##molecule_type mRNA
      ##residues       1-599 ##label WHI
      ##cross-references GB:M38651
COMMENT    The steroid hormones and their receptors are involved in the
             regulation of eukaryotic gene expression and affect cellular
             proliferation and differentiation in target tissues.
COMMENT    In the absence of ligand, steroid hormone receptors are thought to
             be weakly associated with nuclear components; hormone binding
             greatly increases receptor affinity  The hormone-receptor complex
             appears to recognize discrete DNA sequences upstream of
             transcriptional start sites
CLASSIFICATION   #superfamily estrogen receptor; erbA transforming protein
                   homology
KEYWORDS           DNA binding, nucleus, phosphoprotein, steroid binding,
                     steroid hormone receptor; transcription regulation; zinc
                     finger
FEATURE
   1-183                  #domain amino-terminal #label NH2\
   184-275                #domain DNA binding #status predicted #label DNA\
   187-460                #domain erbA transforming protein homology #label ERBA\
   187-210                #region zinc finger CCCC motif\
   223-247                #region zinc finger CCCC motif\
   260-275                #region nuclear location signal\
   304-556                #domain steroid binding #status predicted #label STB\
   189,192,206,209        #binding_site zinc (Cys) #status predicted\
   225,231,241,244        #binding_site zinc (Cys) #status predicted\
   240,309                #binding_site phosphate (Ser) (covalent) #status
                            predicted
SUMMARY             #length 599  #molecular-weight 66154  #checksum 4743

  Query Match             46.8%;  Score 1455;  DB 3;  Length 599;
  Best Local Similarity   51.7%;  Pred. No. 4.55e-252;
  Matches   214;  Conservative   93;  Mismatches   87;  Indels   20;  Gaps   14;

Db     108 spsplmllhppqlspflhphgqqvpyylenepsayavrdtgppafyrsnsdnrrqngr 166
           : || :|    | :|||  |   | |      | | : |:      |            :: :|: :|
Qy      20 TTSFNVLWPTPGHLSF LVVHRQLSHLYASFQKSFWCEARSLEHILPVNRELLKRKVSG- 77

Db     167 erlsssnekgnmimesakettyaavcndyasgyhygvwsceackaffkrsiqghndymcp 226
           :|  :|    |       | :::::||'|:||||||||||||||||||||||||||||:||
Qy      78 NR-CASPVTGP---GSKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICP 133

Db     227 atnqctidknrrksqacrlrkcyevgmmkggirkdrrggrmlkhkrqrddlegrnemga 286
           ||||||||||||||||||||||||||||||:| | |::| | |:::  |||    : :    |
Qy     134 ATNQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVR--RQRSA-DEQLH-CA 189

Db     287 sgdmraanlwpsplvikhtkknspalsltadqmvsalldaeppmiyseydpsrpfseasm 346
           :  :|::      :| | :       :  |||  ::|:| :||:|||| :       || ||:|:||
Qy     190 GKAKRSGG--HAPRV-RELLLD--ALS--PEQLVLTLLEAEPPHVLIS-RPSAPFTEASM 241

Db     347 mglltnladrelvhminwakrvpgfgdlnlhdqvhllecawleilmiglvwrsmehpqkl 406
           |  ||:|||:||||||:|||::||| :|:| |||:|||  |:|:||:||:|||::|||||
Qy     242 MMSLTKLADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKL 301

Db     407 lfapnllldrnqgkcveqmveifdmllatssrfrmmnlqgeefvclksiillnsgvytfl 466
           :|||:|:||::||||||::||||||||||||:||||  ::||  |::|:|::|||||::| ::
Qy     302 IFAPDLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLV 361

Db     467 sstlksleekdhihrvldkitdtlihlmakagltlqqqhrrlaqllliishirh 520
           ::|  : :     : ::|: :||:|: ::||:|:: |||  ||| ||::|||:||
Qy     362 IATQDA-DSSRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRH 414
```

```
RESULT     6
ENTRY              QRRTE      #type complete
TITLE              estrogen receptor - rat
ORGANISM           #formal_name Rattus norvegicus #common_name Norway rat
DATE               30-Sep-1991 #sequence_revision 14-Jul-1994 #text_change
                     15-Jun-1996
ACCESSIONS         S07379; S16731
REFERENCE          S07379
   #authors        Koike, S.; Sakai, M.; Muramatsu, M.
   #journal        Nucleic Acids Res. (1987) 15:2499-2513
   #title          Molecular cloning and characterization of rat estrogen
                     receptor cDNA.
   #cross-references MUID:87174780
   #accession      S07379
      ##molecule_type mRNA
      ##residues       1-600 ##label KOI
      ##cross-references EMBL:Y00102
REFERENCE          S16731
   #authors        Maggi, A.M.A.
   #submission     submitted to the EMBL Data Library  June 1991
   #accession      S16731
      ##molecule_type mRNA
      ##residues       1-487,'T',489-600 ##label MAG
      ##cross-references EMBL:X61098
COMMENT    The steroid hormones and their receptors are involved in the
             regulation of eukaryotic gene expression and affect cellular
             proliferation and differentiation in target tissues.
COMMENT    In the absence of ligand, steroid hormone receptors are thought to
             be weakly associated with nuclear components; hormone binding
             greatly increases receptor affinity. The hormone-receptor complex
             appears to recognize discrete DNA sequences upstream of
             transcriptional start sites.
CLASSIFICATION   #superfamily estrogen receptor; erbA transforming protein
                   homology
KEYWORDS           DNA binding; nucleus; phosphoprotein; steroid binding;
                     steroid hormone receptor; transcription regulation; zinc
                     finger
FEATURE
   1-184                  #domain amino-terminal #label NH2\
   185-276                #domain DNA binding #status predicted #label DNA\
   188-461                #domain erbA transforming protein homology #label ERBA\
   188-211                #region zinc finger CCCC motif\
   224-246                #region zinc finger CCCC motif\
   261-276                #region nuclear location signal\
   305-557                #domain steroid binding #status predicted #label STB\
   190,193,207,210        #binding_site zinc (Cys) #status predicted\
   226,232,242,245        #binding_site zinc (Cys) #status predicted\
   241,310                #binding_site phosphate (Ser) (covalent) #status
                            predicted
SUMMARY             #length 600  #molecular-weight 67029  #checksum 4599

  Query Match             46.6%,  Score 1448;  DB 3;  Length 600;
  Best Local Similarity  51.0%;  Pred. No. 1.10e-250;
  Matches  211;  Conservative   97;  Mismatches  86;  Indels  20;  Gaps  14;

Db     109 spsplmllhppphvspflhphghqvpyylenepsay-avrdtgppafyrsnsdnrrqngr 167
           : || :|    | |:|| |  |  :    | |  : |:    |           :: :|: :|
Qy      20 TTSPNVLWPTPGHISP-LVVHRQLSHLYAEPQKSPWCEARSLEHTLPVNRETLKRKVSG- 77

Db     168 erlssssekgnmimesaketrycavcndyasgyhygvwscegckaffkrsiqghndymcp 227
           :|  :|:  |       | :::::||||:||||||||||||||||||||||||||||:|,
Qy      78 NR-CASPVTGP---GSKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICP 133

Db     228 atnqctidknrrksqacrlrkcyevgmmkggirkdrrggrmlkhkrqrddlegrnemgt 287
           ||||||||||||||||||||||||||||:| | |::| | |:::: |  |: : :    |
Qy     134 ATNQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQRSADE-QLHCA-GK 191

Db     288 sqdmraanlwpsplvikhtkknspalsltadqmvsalldaepplivseydpsrpfseasm 347
           :  :|::     :| | :     :   |||  ::|:| :||:|||| :       || ||:||||
Qy     192 A--KRSGG--HAPRV-RELLLD--ALS--PEQLVLTLLEAEPPHVLISRPSAPFTEASM 241

       348 mglltnladrelvhminwakrvpgfgdlnlhdqvhllecawleilmiglvwrsmehpgkl 407
           |  ||:|||:|||||||:|||::|||  :|:| |||:|||  |:|:||:||:|||::|||||
Qy     242 MMSLTKLADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKL 301

Db     408 lfapnllldrnqgkcvegmveifdmllatssrfrmmnlqgeefvclksiillnsgvytfl 467
           :|||:|:|||::||||||::||||||||||:||||  :::||  |::|::|:||||||::| ::
Qy     302 IFAPDLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLV 361

Db     468 sstlksleekdhihrvldkindtlihlmakagltlqqqhrrlaqllllilshirh 521
           ::|  :  :    : ::|: : |:|: ::||:|:: |||  ||| ||::|||:||
Qy     362 TATQDA-DSSRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRH 414


RESULT     7
ENTRY              A37197     #type complete
TITLE              estrogen receptor - rainbow trout
ORGANISM           #formal_name Oncorhynchus mykiss #common_name rainbow trout
DATE               31-Jan-1992 #sequence_revision 31-Jan-1992 #text_change
                     03-Jun-1996
ACCESSIONS         A37197; A40070
REFERENCE          A37197
   #authors        Pakdel, F.; Le Gac, F.; Le Goff, P.; Valotaire, Y.
   #journal        Mol. Cell. Endocrinol. (1990) 71:195-204
   #title          Full-length sequence and in vitro expression of rainbow trout
                     estrogen receptor cDNA.
   #cross-references MUID:91006824
   #accession      A37197
      ##status          preliminary
      ##molecule_type mRNA
      ##residues        1-574 ##label PAK
      ##cross-references GB:M31559
REFERENCE          A40070
   #authors        Pakdel, F.; Le Guellec, C.; Vaillant, C.; Le Roux, M.G.;
                     Valotaire, Y.
   #journal        Mol. Endocrinol. (1989) 3:44-51
   #title          Identification and estrogen induction of two estrogen
                     receptors (ER) messenger ribonucleic acids in the rainbow
                     trout liver: sequence homology with other ERs.
   #cross-references MUID:89127284
   #accession      A40070
      ##status          preliminary
      ##molecule_type mRNA
      ##residues        150-574 ##label PA2
      ##cross-references GB:M31559
CLASSIFICATION     #superfamily estrogen receptor; erbA transforming protein
                     homology
KEYWORDS           DNA binding; nucleus; steroid hormone receptor; zinc finger
FEATURE
   145-421                #domain erbA transforming protein homology #label ERBA\
   147-167                #region zinc finger\
   183-207                #region zinc finger
SUMMARY            #length 574  #molecular-weight 62871  #checksum 1057

  Query Match             42.7%;  Score 1327;  DB 7;  Length 574;
```

```
  Best Local Similarity  57.5%;  Pred. No. 7.87e-227;
  Matches  196;  Conservative   64;  Mismatches  61;  Indels  20;  Gaps  15;

Db     143 etrycavcsdfasgyhygvwscegckaffkrsiqghndymcpatnqctmdrnrrksgqac 202
           ::..||||||:||||||||||||||||||||||||||||.||||||||:|:||||||||||
Qy      92 DAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQAC 151

Db     203 rlrkcyevgmvkgglrkdrgg-rvlrkdkryggpagdrekpytvwstgqrpqdgarnsss 261
           ||||||||||||| | |::| | |::|:: |    :| | |     : : : |:  ||:  ::
Qy     152 RLRKCYEVGMVKCGSRRERCGYRLVRRQ-R---SA-D-EQLHCA-GKAKRS--GGH--AP 200

Db     262 slnggggwrqpritmppeqvlfllqgqtp-alcsrqkvarpytevtmmtlltsmadkelv 320
             :          : :: | : || ||:::  |   | ||    | |:|| :||  || :||||||
Qy     201 RVREL--LLDA-LS-PEQLVLTLLEAEPPHVLISRPS-A-PFTEASMMMSLTKLADKELV 254

Db     321 hmiawakkvpgfqelslhdqvqllesswlevlmigliwrsihcpgklifaqdlildrseq 380
           |||:|||||:||| |||| |||:|||| |:||||:||:||||   ||||||| ||:||| ||
Qy     255 HMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEG 314

Db     381 dcvegmaeifdmllatvsrfqmlklkpeefvclkailllnpgafsfcsnsveslhnssav 440
            ||||: ||||||||| ||| |||    |..|:|||:||||:: ::: : : :: :| :
Qy     315 KCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSRKL 373

Db     441 esmldnitdalihhishsgasvqqqprrqaqlllllshirh 481
            .|. :||||. |. || | |||. | | ||.||||.||
Qy     374 AHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRH 414


RESULT     8
ENTRY              S58224     #type fragment
TITLE              oestrogen receptor - Atlantic salmon (fragment)
ORGANISM           #formal_name Salmo salar #common_name Atlantic salmon
DATE               13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change
                     26-Apr-1996
ACCESSIONS         S58224
REFERENCE          S58224
   #authors        Rogers, S.A.; Llewellyn, L.; Ramsurn, V.P.; Sweeney, G.E.;
                     Wigham, T.
   #submission     submitted to the EMBL Data Library, July 1995
   #accession      S58224
      ##status          preliminary
      ##molecule_type mRNA
      ##residues        1-535 ##label ROG
      ##cross-references EMBL:X89959
CLASSIFICATION     #superfamily erbA transforming protein homology
FEATURE
   103-382                #domain erbA transforming protein homology #label ERBA
SUMMARY            #length 535  #checksum 9963

  Query Match             42.3%;  Score 1314;  DB 7;  Length 535;
  Best Local Similarity  56.9%;  Pred. No. 2.85e-224;
  Matches  195;  Conservative   61;  Mismatches  66;  Indels  21;  Gaps  14;

Db     101 etrycavcsdfasgyhygfwscegckaffkrsiqghndymcpatnqctmdrnrrksgqac 160
           ::::||||||:||||||| ||||||||||||||||||||:|||||||||:|:||||||||||
Qy      92 DAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQAC 151

Db     161 rlrkcyevgmvkgglrkdrgg-rvlrkdkryggpagdrekpyqdlehrtappqdgarnss 219
           ||||||||||||| | |::| | |::|:: |    :| : :   :     : | : ||: :
Qy     152 RLRKCYEVGMVKCGSRRERCGYRLVRRQ-R---SADE-QLHCAG---K-AK-RSGGHAPR 201

Db     220 ssslsggggwrqpritmppeqvlfllqgaeppalcsrqkvarpytevtmmtlltsmadke 279
               |          : :: | : || ||::  |   | ||    | |:|| :||  || : :||
Qy     202 VRELL----L-DA-LS-PEQLVLTLLEAEPPHVLISRPS-A-PFTEASMMMSLTKLADKE 252

Db     280 lvhmiawakkvpgfqelslhdqvqllesswlevlmigliwrsihcpgklifaqdlildrs 339
           |||||:||||:||| |||| |||:|||| |:||||:||:||||   ||||||| ||:||
Qy     253 LVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRD 312

Db     340 eqdcvegmaeifdmllatvsrfrmlklkpeefvclkailllnsgafsfcsnsveslhnss 399
           || ||||: ||||||||| ||| |||    |::|:|||:||||:: ::: : : :: :
Qy     313 EGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSR 371
```

```
Db    400 avesmldnitdalihhishsqasvqqqprrqvqlllllshirh 442
            :  :|: :||||:  |: || | |||: |    ||:||||:||
Qy    372 KLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRH 414

RESULT     9
ENTRY               S58087     #type complete
TITLE               estrogen receptor related protein - mouse
ORGANISM            #formal_name Mus musculus #common_name house mouse
DATE                13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change
                      26-Apr-1996
ACCESSIONS          S58087
REFERENCE           S58087
   #authors         Pettersson, K.; Svensson, K.; Mattsson, R.; Carlsson, B.;
                      Ohlsson, R.; Berkenstam, A.
   #submission      submitted to the EMBL Data Library, July 1995
   #description     Expression of a novel member of estrogen response
                      element-binding nuclear receptors is restricted to the
                      early stages of chorion formation during mouse
                      embryogenesis.
   #accession       S58087
      ##status          preliminary
      ##molecule_type mRNA
      ##residues        1-433 ##label PET
      ##cross-references EMBL:X89594
CLASSIFICATION    #superfamily erbA transforming protein homology
FEATURE
   101-352              #domain erbA transforming protein homology #label ERBA
SUMMARY             #length 433  #molecular-weight 48400  #checksum 1075

  Query Match              23.8%;  Score 740;  DB 14;  Length 433;
  Best Local Similarity  42.5%;  Pred. No. 1.41e-112;
  Matches   113;  Conservative   57;  Mismatches  85;  Indels  11;  Gaps  10;

Db    101 rlclvcgdiasgyhygvasceackaffkrtiqgnieyncpatneceitkrrrksgqacrf 160
            ::| ||:| |||||||| |||:||||||||:|||: :| ||||||:| | | |||||||||:
Qy     94 HFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRL 153

Db    161 mkclkvgmlkegvrldrv--r-qqrqky-krrldsenspylnlpisppakkpltkivs-n 215
            ||  |||:| | | :|    |    ||:      :|         :  :    :|  :  |   :| :
Qy    154 RKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPE 213

Db    216 -l-lqv-eqdklyam ppndipeqdikalttlceladrelvflinwakhipgfpsltlgd 271
            | |  : |  :      :  :     |   :      : :|  |||:|||  :|:||| ||||  |:| |
Qy    214 QLVLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 273

Db    272 qmsllqsawmeilllglvyrslpyddklayaedyimdeehsrlv-qlldlyrallqlvrr 330
            |: ||:| |||:|::|:: ||:      || :| | ::|  :  ::  | |:|:::   :|    |
Qy    274 QVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSR 333

Db    331 ykklkvekeefmilkalalansdsmy 356
            :: ||::   |:: :||: | || |||
Qy    334 FRELKLQHKEYLCVKAMILLNS SMY 358

RESULT    10
ENTRY               B29345     #type complete
TITLE               steroid hormone receptor ERR2 precursor - human
ORGANISM            #formal_name Homo sapiens #common_name man
DATE                15-Dec-1988 #sequence_revision 15-Dec-1988 #text_change
                      03-Jun-1996
ACCESSIONS          B29345
REFERENCE           A29345
   #authors         Giguere, V.; Yang, N.; Segui, P.; Evans, R.M.
   #journal         Nature (1988) 331:91-94
   #title           Identification of a new class of steroid hormone receptors.
   #cross-references MUID:88122546
   #accession       B29345
      ##molecule_type mRNA
      ##residues        1-433 ##label GIG
      ##cross-references GB:X51417; GB:Y00290
CLASSIFICATION    #superfamily erbA transforming protein homology
KEYWORDS            DNA binding; zinc finger
FEATURE
   101-352              #domain erbA transforming protein homology #label ERBA\
   103-123              #region zinc finger\
   139-163              #region zinc finger
SUMMARY             #length 433  #molecular-weight 48287  #checksum 1133

  Query Match              23.7%;  Score 736;  DB 13;  Length 433;
  Best Local Similarity  43.6%;  Pred. No. 8.21e-112;
  Matches   116;  Conservative   55;  Mismatches  84;  Indels  11;  Gaps   9;

Db    101 rlclvcgdiasgyhygvasceackaffkrtiqgnieyscpatneceitkrrrksgqacrf 160
            ::| ||:| |||||||| |||:||||||||:|||: :| ||||||:| | | |||||||||:
Qy     94 HFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRL 153

Db    161 mkclkvgmlkegvrldrv--r-qgrqky-krrldsenspylslqlsppakkpltkivs-  214
            ||  |||:| | | :|    |    ||:      :|        :  |    :|  :  |    :|
Qy    154 RKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPE 213

Db    215 yl-l-vaepdklyam-ppddvpeqdikalttlcdladrelvflinwakhipgfsnltlqd 271
            | | . |..     . .    |  .    . .| |||:||| :||||| |||| :|:| |
Qy    214 QLVLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 273

Db    272 qmsllqsawmeilllglvyrslpyddklayaedyimdeehsrlv-qllelyrallqlvrr 330
            |: ||:| |||:|::|:: ||:     || :| | ::|  : ::  | |:||::   :|     |
Qy    274 QVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSR 333

Db    331 ykklkvekeefvmlkalalansdsmy 356
            :: ||::   |:: :||: | || |||
Qy    334 FRELKLQHKEYLCVKAMILLNS-SMY 358

RESULT    11
ENTRY               A29345     #type complete
TITLE               steroid hormone receptor ERR1 precursor - human
ALTERNATE_NAMES   estrogen-related receptor
ORGANISM            #formal_name Homo sapiens #common_name man
DATE                15-Dec-1988 #sequence_revision 15-Dec-1988 #text_change
                      03-Jun-1996
ACCESSIONS          A29345; A49074
REFERENCE           A29345
   #authors         Giguere, V.; Yang, N.; Segui, P.; Evans, R.M.
   #journal         Nature (1988) 331:91-94
   #title           Identification of a new class of steroid hormone receptors.
   #cross-references MUID:88122546
   #accession       A29345
      ##molecule_type mRNA
      ##residues        1-521 ##label GIG
      ##cross-references EMBL:X51416; EMBL:Y00290
REFERENCE           A49074
   #authors         Wiley, S.R.; Kraus, R.J.; Zuo, F.; Murray, E.E.; Loritz, K.;
                      Mertz, J.E.
   #journal         Genes Dev. (1993) 7:2206-2219
   #title           SV40 early-to-late switch involves titration of cellular
                      transcriptional repressors.
   #accession       A49074
      ##status          preliminary
      ##molecule_type protein
      ##residues        166-169,'X',171-173 ##label WIL
CLASSIFICATION    #superfamily erbA transforming protein homology
KEYWORDS            DNA binding; nucleus; zinc finger
FEATURE
   174-434              #domain erbA transforming protein homology #label ERBA\
   176-196              #region zinc finger\
   212-236              #region zinc finger
SUMMARY             #length 521  #molecular-weight 55742  #checksum 5840

  Query Match              23.3%;  Score 725;  DB 13;  Length 521;
  Best Local Similarity  40.1%;  Pred. No. 1.04e-109;
  Matches   130;  Conservative   67;  Mismatches 117;  Indels  10;  Gaps   7;

Db    174 rlclvcgdvasgyhygvasceackaffkrtiqgsieyscpasneceitkrrrkacqacrf 233
            ::| ||:| |||||||| |||:||||||||:|||  :| |||:|:| | | |||:||||||:
Qy     94 HFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRL 153
```

```
Db    234 tkqlrvgmlkegvrldrvrggrqkykrrpevdplpfpg-p-fpagplavaggprktaapv 291
           ||   |||:| | | :|  | |    :.|.  ' |   |  .    |         | .
Qy    154 PKCYEVGMVKCGSPPEPC-GYPLVPPQPSADEQLHCAGKAKPSGGHAPPVPELLLDALSP 212

Db    292 nalvshllvvepeklyampdpagpdghlpavatlcdlfdreivvtiswaksipgfsslsl 351
          : ||  ||  ||  :      |::|      |  |  |:|:|  ||||| ||||  |||
Qy    213 EQLVLTLLEAEPPHVL-ISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSL 271

Db    352 sdqmsvlqsvwmevlvlgvaqrslplqdclafaedlvld-eegaraaglgelgaallqlv 410
           ||: :|:| |||||::|:  ||     | || |||||  ||     |· |·    ||
Qy    272 FDQVPLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATT 331

Db    411 rrlqalrlereeyvllkalalansdsvhiedeprlwsscekllh--cal---leyeagra 465
           |:: |:|:: ||: :||: | ||     :     :    |  || |  :|:    |    :::
Qy    332 SRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKS 391

Db    466 gpgggaerrragrllltlpllrqt 489
          | :.  .  | : ||  |   | .
      392 GISSQQQSMRLANLLMLLSHVRHA 415

RESULT   12
ENTRY              A39596     #type complete
TITLE              progesterone receptor B form - mouse
ORGANISM           #formal_name Mus musculus #common_name house mouse
DATE               20-Mar-1992 #sequence_revision 20-Mar-1992 #text_change
                     03-Jun-1996
ACCESSIONS         A39596
REFERENCE          A39596
   #authors        Schott, D.R.; Shyamala, G.; Schneider, W.; Parry, G.
   #journal        Biochemistry (1991) 30:7014-7020
   #title          Molecular cloning, sequence analyses, and expression of
                     complementary DNA encoding murine progesterone receptor.
   #cross-references MUID:91299759
   #accession      A39596
      ##status           preliminary
      ##molecule_type mRNA
      ##residues         1-923 ##label SCH
      ##cross-references GB:J05333
CLASSIFICATION     #superfamily progesterone receptor, erbA transforming protein
                     homology
KEYWORDS           DNA binding; nucleus; steroid hormone receptor; zinc finger
FEATURE
   555-919               #domain erbA transforming protein homology #label ERBA\
   557-577               #region zinc finger\
   593-617               #region zinc finger
MMARY              #length 923  #molecular-weight 99073  #checksum 6627

 Query Match             21.9%;  Score 681;  DB 7;  Length 923,
 Best Local Similarity   30.8%;  Pred. No. 2.56e-101;
 Matches  127,  Conservative  112,  Mismatches 138,  Indels  27,  Gaps  27,

Db    491 aapgtaaapaiyqplg-lnglpqlg-y-qaavlkdslpqvyppylnylrpdseasqspqy 547
          :::|   :      |  |   ||      |    |     || |     |     :       ||
Qy     28 PTPGHLSPLVVHPQLSHLYAEPQKSPWCEAPSLEHTLP-VNRETLK-RKVSGNPCASPVT 95

Db    548 gfdslpq-kiclicgdeasgchygvltcgsckvffkrameggqhnylcagrndcivdkirr 606
          |  |  :   :| :|·| ||| ||||   |  ||  |||||·· .|  :.:|  |·  |:|   || ||
Qy     86 GPGSKPLAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRR 145

Db    607 kncpacrlrkccqagmvl-ggrkfk-kfnkvrvmrtldgvalpqsvglpnesqalsqrit 664
          |:| |||||||  : |||   |:|:  :   ·  ||  |   |   | :   |  · · |  · · |·
Qy    146 KSCQACRLRKCYEVGMVKCGSPPEPCGYPLVPPQPSADE-QL-HCAGKAKPSGGHAPPVP 203

Db    665 fspnqeiqlvp-plinllmsiepdviyaghdntkpdtsssllltslnqlgerqllsvvkws 723
              ·   |  |  |·  |    ||   ·      ·   |  |  |:   ||   |    ·|:  ·  |·
Qy    204 -ELLLDA-LSPEQLVLTLLEAEPPHVLISRPSA-PFTEASMMMSLTKLADKELVHMISWA 260

Db    724 kslpgfrnlhiddqitliqyswmslmvfglgwrsykhvsggmlyfapdliln-eq-r-mk 780
          | :|||  :|  :  ||:  |::   ||    .   || |||    |    |  .   |||||·|·  ::  ·  ·
Qy    261 KKIPGFVELSLFDQVPLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVE 318
```

```
Db    781 e-lsfyslcltmwqipqefvklqvtheeflcmkvllllnt-i-pleglrsqsqfeemrss 837
            | :. . |            |   |  · | | ||·|  :||||.   ||          .    : :
Qy    319 GILEIFDM-L-L-ATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVIATQDADSSRKLAH 375

Db    838 yirelikaig-lrqkgvvptsqrfyqltklldslhdlvkq 876
           .       |·      |      .   |.    :|..||   |    |··
Qy    376 LLNAVIDALVWVIAKSGISSQQQSMRLANLL-MLLSHVRH 414

RESULT   13
ENTRY              I53280     #type complete
TITLE              progesterone receptor - rat
ORGANISM           #formal_name Rattus norvegicus #common_name Norway rat
DATE               02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change
                     02-Aug-1996
ACCESSIONS         I53280
REFERENCE          I53280
   #authors        Park-Sarge, O.K.; Mayo, K.E.
   #journal        Endocrinology (1994) 134:709-718
   #title          Regulation of the progesterone receptor gene by gonadotropins
                     and cyclic adenosine 3',5'-monophosphate in rat granulosa
                     cells.
   #cross-references MUID:94130817
   #accession      I53280
      ##status           preliminary; translated from GB/EMBL/DDBJ
      ##molecule_type mRNA
      ##residues         1-923 ##label RES
      ##cross-references GB:L16922; NID:g463282; CDS_PID:g463283
SUMMARY            #length 923  #molecular-weight 99407  #checksum 9843

 Query Match             21.6%;  Score 671;  DB 14;  Length 923;
 Best Local Similarity   32.6%;  Pred. No. 2.04e-99;
 Matches  108,  Conservative   94,  Mismatches 108,  Indels  21,  Gaps  21;

Db    556 iclicgdeasgchygvltcgsckvffkrameggqhnylcagrndcivdkirrkncpacrlr 615
          :| .| | ||| |||| |  || ||||···|···|·|:· |:| :|| |||:| |||||
Qy     95 FCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLR 154

Db    616 kccqagmvl-ggrkfk-kfnkvrvmraldgvalpqsvafpnesqtlgqritfspnqeiql 673
          || : ||| |:|· ·  · || |: |  |           |  : |        . |
Qy    155 KCYEVGMVKCGSPPEPCGYPLVPPQPSADE-QL-HCAGKAKPSGGHAPPVP-ELLLDA-L 210

Db    674 vp-plinllmsiepdvvyaghdntkpdtsssllltslnqlgerqllsvvkwskslpgfrnl 732
           |  |   |.  ||  |  .:    | |  |·· ||  |····|· ·· |:| :||| :|
Qy    211 SPEQLVLTLLEAEPPHVLISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVEL 269

Db    733 hiddqitliqyswmslmvfglgwrsykhvsgqmlyfapdliln-eq-r-mke-lsfyslc 788
           : ||· |    || :   || |||  | ·|   · ||||| |  .. .  .   | .. .
Qy    270 SLFDQVPLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDM- 326

Db    789 ltmwqipqefvklqvtheeflcmkvllllnt-i-pleglrsqsqfeemrssyirelikal 846
          |  ·     |  |  .  |  |  ||·|  · ||!·   ||     . ·   :  ·  ·  ·  ·  |·
Qy    327 I-I-ATTSPFPELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVIDAL 384

Db    847 g-lrqkgvvpssqrfyqltklldslhdlvkq 876
           .  |: ..| |:  .|..||  |   |:.
Qy    385 VWVIAKSGISSQQQSMRLANLL-MLLSHVRH 414

RESULT   14
ENTRY              QRHUP      #type complete
TITLE              progesterone receptor form B - human
ALTERNATE_NAMES    hPR
CONTAINS           progesterone receptor form A
ORGANISM           #formal_name Homo sapiens #common_name man
DATE               30-Jun-1987 #sequence_revision 18-Nov-1994 #text_change
                     15-Nov-1996
ACCESSIONS         S09971, S12464; A03245
REFERENCE          S09971
   #authors        Kastner, P.; Krust, A.; Turcotte, B.; Stropp, U.; Tora, L.;
                     Gronemeyer, H.; Chambon, P.
   #journal        EMBO J. (1990) 9:1603-1614
   #title          Two distinct estrogen-regulated promoters generate
```

```
                   transcripts encoding the two functionally different human
                     progesterone receptor forms A and B
   #cross-references MUID:90228361
   #accession      S09971
      ##molecule_type mRNA
      ##residues       1-933 ##label KAS
      ##cross-references EMBL:X51730
REFERENCE          S12464
   #authors        Kastner, P.
   #submission     submitted to the EMBL Data Library, February 1990
   #accession      S12464
      ##molecule_type mRNA
      ##residues       1-343,'T',345-933 ##label KA2
      ##cross-references EMBL:X51730
REFERENCE          A03245
   #authors        Misrahi, M.; Atger, M.; D'Auriol, L.; Loosfelt, H.; Meriel,
                     C.; Fridlansky, F.; Guiochon-Mantel, A.; Galibert, F.;
                     Milgrom, E.
   #journal        Biochem. Biophys. Res. Commun. (1987) 143:740-748
   #title          Complete amino acid sequence of the human progesterone
                     receptor deduced from cloned cDNA.
   #cross-references MUID:87184565
   #accession      A03245
      ##molecule_type mRNA
      ##residues       1-225,'G',227-255,'V',257-659,'V',661-933 ##label MIS
      ##cross-references GB:M15716
GENETICS
   #gene           GDB:PGR
      ##cross-references GDB:119493
   #map_position 11q22.1-11q22.3
CLASSIFICATION   #superfamily progesterone receptor; erbA transforming protein
                   homology
KEYWORDS         alternative splicing; DNA binding; glycoprotein; nucleus;
                   phosphoprotein; steroid hormone receptor; transcription
                   regulation; zinc finger
FEATURE
   1-933                #product progesterone receptor form B #status predicted
                          #label MA1\
   165-933              #product progesterone receptor form A #status predicted
                          #label MA2\
   565-829              #domain erbA transforming protein homology #label ERBA\
   567-587              #region zinc finger CCCC motif\
   603-627              #region zinc finger CCCC motif\
   681-933              #domain steroid binding #status predicted #label STE\
   41                   #binding_site phosphate (Thr) (covalent) #status
                          predicted\
   227,232,552,793      #binding_site phosphate (Ser) (covalent) #status
                          predicted\
   329,374,601          #binding_site phosphate (Tyr) (covalent) #status
                          predicted\
   664                  #binding_site carbohydrate (Asn) (covalent) #status
                          predicted
SUMMARY            #length 933  #molecular-weight 99028  #checksum 5865

  Query Match             21.4%;  Score 666;  DB 3;  Length 933;
  Best Local Similarity   33.4%;  Pred. No. 1.82e-98;
  Matches  110;  Conservative   89;  Mismatches 113;  Indels   17;  Gaps   17;

Db     566 iclicgdeasgchygvltcgsckvffkramegqhnylcagrndcivdkirrkncpacrlr 625
            :| :|:| ||| |||| :| :|| ||||:::|:::|:|:: |:| :|| |||:| |||||
Qy      95 FCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRLR 154

Db     626 kccqagmvl-ggrkfk-kfnkvrvvraldavalpqplgvpnesqalsqrftfspgqdiql 683
            || : |||  |:|: :   :  ||  |: |     | :  | :: | : : |        | |
Qy     155 KCYEVGMVKGGSRRERCGYRLVRRQRSADE-QL-HCAGKAKRSGGHAPRVR-ELLLDA-L 210

Db     684 ip-plinllmsiepdviyaghdntkpdtssslltslnqlgerqllsvvkwskslpgfrnl 742
             |  |:  |:  ||  :  :: :: | | :|:: ||  |::::|: :: |:| :||| :|
Qy     211 SPEQLVLTLLEAEPPHVLISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVEL 269

Db     743 hiddqitliqyswmslmvfglgwrsykhvsgqmlyfapdlilneqrmk-essfyslcltm 801
            : ||: |::  || ::: || |||  | :| :: |||||:|. .  |   ..  .   :
Qy     270 SLFDQVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDML 327
```

```
Db     802 wqipqefvklqvsqeeflcmkvllllntiplegirsqtq-feemrs-syi-relikaiq- 857
                 |  | : : |:||:| ::|||: ::  | : ||  :  |  : :    : |:
Qy     328 LATTSRFRELKLQHKEYLCVKAMILLNS-SMYPLVTATQDADSSRKLAHLLNAVTDALVW 386

Db     858 lrqkgvvsssqrfyqltklldnlhdlvkq 886
            :  |: :|| |:  :|::||  |   |::
Qy     387 VIAKSGISSQQQSMRLANLL-MLLSHVRH 414


RESULT   15
ENTRY              A25923      #type complete
TITLE              progesterone receptor - rabbit
ORGANISM           #formal_name Oryctolagus cuniculus #common_name domestic
                     rabbit
DATE               16-Aug-1988 #sequence_revision 16-Aug-1988 #text_change
                     03-Jun-1996
ACCESSIONS         A25923
REFERENCE          A25923
   #authors        Loosfelt, H.; Atger, M.; Misrahi, M.; Guiochon-Mantel, A.;
                     Meriel, C.; Logeat, F.; Benarous, R.; Milgrom, E.
   #journal        Proc. Natl. Acad. Sci. U.S.A. (1986) 83:9045-9049
   #title          Cloning and sequence analysis of rabbit progesterone receptor
                     complementary DNA.
   #cross-references MUID:87067449
   #accession      A25923
      ##molecule_type mRNA
      ##residues       1-930 ##label LOO
CLASSIFICATION   #superfamily progesterone receptor; erbA transforming protein
                   homology
KEYWORDS         DNA binding; nucleus; steroid hormone receptor; zinc finger
FEATURE
   566-826              #domain erbA transforming protein homology #label ERBA\
   568-588              #region zinc finger\
   604-628              #region zinc finger
SUMMARY            #length 930  #molecular weight 98666  #checksum 2858

  Query Match             21.3%;  Score 662;  DB 7;  Length 930;
  Best Local Similarity   32.5%;  Pred. No. 1.05e-97;
  Matches  106;  Conservative   91;  Mismatches 114;  Indels   15;  Gaps   15;

Db     567 iclicgdeasgchygvltcgsckvffkramegqhnylcagrndcivdklrrkncpacrlr 626
            .| .|.| ||| |||| .| .|| |||:..|...|.|:: |:| :|| |||:| |||||
Qy      95 FCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRLR 154

Db     627 kccqagmvl-ggrkfk-kfnkvrvmraldavalpqpvgipnesq-ritfspsqelq-lip 682
            || : |||  |:|: :   :  ||  |: |     | :  | :: | :           :: | |
Qy     155 KCYEVGMVKGGSRRERCGYRLVRRQRSADE-QL-HCAGKAKRSGGHAPRVKELLLDALSP 212

Db     683 -plinllmsiepdviyaghdntkpdtsssllts lnqlgerqllsvvkwskslpgfrnlhi 741
             |.  |.  ||  .  .. .. | | .|.. ||  |::::|: :: |:| :||| :| :
Qy     213 EQLVLTLLEAEPPHVLISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSL 271

Db     742 ddqitliqyswmslmvfglgwrsykhvsgqmlyfapdlilneqrmk-essfyslcltmwq 800
            ||: |::  || ::: || |||  | :| :: |||||:|: :  |   ::  :   :
Qy     272 FDQVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDMLLA 329

Db     801 ipqefvklqvsqeeflcmkvllllnt-i-pleglrsqsqfeemrssyirelikaiq-lrq 857
                 |  | : : |:||:| ::|||: : ||     :::   : : : : |: :
Qy     330 TTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIA 389

Db     858 kqvvsssqrfyqltklldnlhdlvkq 883
           |: :|| |:  :|::||  |   |::
Qy     390 KSGISSQQQSMRLANLL-MLLSHVRH 414


Search completed. Thu Dec 18 08.56.10 1997
Job time : 117 secs.
```

```
***********************************************************************
                         MPSRCH (TM)
***********************************************************************

          Release 2.1D John F. Collins, Biocomputing Research Unit
          Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
                 Distribution rights by IntelliGenetics, Inc.

MPsrch_pp   protein - protein database search, using Smith-Waterman algorithm

 on:             Thu Dec 18 08:57:29 1997   MasPar time 7.79 Seconds
                                            351.052 Million cell updates/sec
ular output not generated

Title:            >US-08-826-361A-6
Description:      (1-416) from US08826361A.pep
Perfect Score:    3108
Sequence:         1 MNYSIPSNVTNLEGGPGRQT..........QQSMRLANLLMLLSHVRHAR 416

Scoring table:    PAM 150
                  Gap 11

Searched:         54101 seqs, 4836506 residues

Post-processing:  Minimum Match 0%
                  Listing first 45 summaries

Database:         a-issued
                    1:backl 2:51 3:52 4:53 5:54 6:55 7:56 8:pct90 9:pct91
                    10:PCT92 11:PCT93 12:PCT94 13:PCT95 14:PCT96

Statistics:       Mean 32.526; Variance 144.261; scale 0.225

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                              SUMMARIES
              %
ult         Query
 No.  Score Match Length DB  ID          Description          Pred. No.
-------------------------------------------------------------------------
   1   1428  45.9    410  1  5223606-5   Patent No. 5223606.   4.60e-123
   2    660  21.2    363  1  5223606-6   Patent No. 5223606.   5.64e-54
   3    609  19.6    356  1  5223606-7   Patent No. 5223606.   3.36e-45
   4    577  18.6    795  3  US-07-716-  Sequence 5, Applicati 3.25e-42
   5    506  16.3    533  5  US-07-952-  Sequence 2, Applicatio 1.28e-35
   6    505  16.2    446  5  US-07-952-  Sequence 4, Applicatio 1.58e-35
   7    493  15.9    462  9  PCT-US91-0  Sequence 2, Applicatio 2.03e-34
   8    494  15.9    467  9  PCT-US91-0  Sequence 4, Applicatio 1.64e-34
   9    461  14.8    463  9  PCT-US91-0  Sequence 6, Applicatio 1.81e-31
  10    420  13.5    440  6  US-08-333-  Sequence 8, Applicatio 1.05e-27
  11    411  13.2    454  1  5260432-2   Patent No. 5260432.   6.95e-27
  12    407  13.1    369  1  5223606-3   Patent No. 5223606.   1.61e-26
  13    407  13.1    448  1  5223606-2   Patent No. 5223606.   1.61e-26
  14    393  12.6    462 10  PCT-US92-0  Sequence 4, Applicatio 3.04e-25
  15    393  12.6    797 10  PCT-US92-0  Sequence 2, Applicatio 3.04e-25
  16    378  12.2    462  1  5171671-2   Patent No. 5171671.   7.01e-24
  17    361  11.6    469 13  PCT-US95-1  Sequence 2, Applicatio 2.43e-22
  18    357  11.5     89  6  US-08-561-  Sequence 3, Applicatio 5.59e-22
  19    357  11.5     89  5  US-08-087-  Sequence 3, Applicatio 5.59e-22
  20    353  11.4    494 13  PCT-US95-1  Sequence 1, Applicatio 1.28e-21
  21    349  11.2     66  1  5217867-1   Patent No. 5217867.   2.94e-21
  22    329  10.6    451 13  PCT-US95-1  Sequence 2, Applicatio 1.95e-19
  23    315  10.1     70  3  US-07-737-  Sequence 3, Applicatio 3.31e-18
  24    310  10.0    461 13  PCT-US95-1  Sequence 2, Applicatio 9.26e-18
  25    310  10.0    461 13  PCT-US95-1  Sequence 2, Applicatio 9.26e-18
  26    310  10.0    461  7  US-08-733-  Sequence 2, Applicati 9.26e-18
  27    310  10.0    461  7  US-08-231-  Sequence 2, Applicatio 9.26e-18
  28    307   9.9    443  7  US-08-342-  Sequence 4, Applicatio 1.71e-17
  29    307   9.9    446 13  PCT-US95-1  Sequence 3, Applicatio 1.71e-17
  30    305   9.8    548  6  US-08-333-  Sequence 2, Applicatio 2.44e-17
  31    304   9.8    513  7  US-08-464-  Sequence 2, Applicatio 3.17e-17
  32    304   9.8    558  6  US-08-333-  Sequence 12, Applicati 3.17e-17
  33    303   9.7    410  1  5438126-2   Patent No. 5438126.   3.89e-17
  34    300   9.7    601  6  US-08-333-  Sequence 14, Applicati 7.19e-17
  35    297   9.6    355  1  5223606-4   Patent No. 5223606.   1.33e-16
  36    299   9.6    460  7  US-08-342-  Sequence 2, Applicatio 9.82e-17
  37    282   9.1    437 12  PCT-US94-0  Sequence 2, Applicatio 2.83e-15
  38    265   8.5    433  6  US-08-333-  Sequence 10, Applicati 8.88e-14
  39    239   7.7     89  5  US-08-087-  Sequence 2, Applicatio 2.01e-11
  40    238   7.7     89  6  US-08-561-  Sequence 2, Applicatio 2.01e-11
  41    191   6.1     60 12  PCT-US94-1  Sequence 1, Applicatio 2.06e-07
  42    169   5.4    367  3  US-07-737-  Sequence 4, Applicatio 1.36e-05
  43    141   4.5    272 13  PCT-US95-0  Sequence 23, Applicati 2.51e-03
  44    141   4.5    272  6  US-08-485-  Sequence 23, Applicati 2.51e-03
  45     89   2.9    427  1  5310667-7   Patent No. 5310667.   1.93e+01

                              ALIGNMENTS

RESULT    1
ID   5223606-5     STANDARD;     PRT;   444 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Patent No. 5223606.
XX
CC   Patent No. 5223606
CC       APPLICANT: BLAUDIN DE THE, HUGHES;MARCHIO, AGNES;TIOLLAIS,
CC   PIERRE;DEJEAN, ANNE
CC       TITLE OF INVENTION: STEROID/THYROID HORMONE RECEPTOR-RELATED
CC   PROTEIN INAPPROPRIATELY EXPRESSED IN HUMAN HEPATOCELLULAR CARCINOMA
CC       NUMBER OF SEQUENCES: 11
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER: US/07/134,130
CC         FILING DATE: 17-DEC-1987
CC       PRIOR APPLICATION DATA:
CC   SEQ ID NO:5:
CC         LENGTH: 410
SQ   SEQUENCE   444 AA; 50020 MW; 1110634 CN;

  Query Match            45.9%;  Score 1428;  DB 1;  Length 410;
  Best Local Similarity  57.7%;  Pred. No. 4.60e-123;
  Matches  131;  Conservative   79;  Mismatches   50;  Indels   12;  Gaps    9;

Db        1 CAVCNDYASGYHYGVWSCEGCKAFFKRSIQGHNDYMCPATNQCTIDKNRRKSCQACRLRK 60
            ||||:||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Qy       96 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 155

Db       61 CYEVGMMKGGIRKDRRGGRMLKHKRQRDDGEGRGEVGSAGDMRAANLWPSPLMIKRSKKN 120
            ||||||:| | |..| | |...  |||    | .    .: . .|..     :| : |
Qy      156 CYEVGMVKCGSRRERCGYRLVR--RQRSADE-QLHCAGKA-KRSGG--HAP-RV-RELL- 206

Db      121 SLALSLTADQMVSALLDAEPPILYSEYDPTRPFSEASMMGLLTNLADRELVHMINWAKRVP 180
             |  .:|...:|:| :||.|||          |. ||.|||||  ||:|||.||||||.|||..|
Qy      207 -ID-ALSFEQLVLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIP 264

Db      181 GFVDLTLHDQVHLLECAWLEILMIGLVWRSMEHPVKLLFAPNLLLDRNQGKCVEGMVEIF 240
            |||.|:| |||.|||  |.|.|||.||..||..|| ||.||||.|.|||..||||||::|||
Qy      265 GFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIF 324

Db      241 DMLLATSSRFRMMNLQGEEFVCLKSIILLNSGVYTFLSSTLKSLEEKDHIHRVLDKITDT 300
            ||||||:||||  ..||  |..|.|..||||||..| ....|   . :      : :|: :||:
Qy      325 DMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSRKLAHLLNAVTDA 383
```

```
Db     301 LIHLMAKAGLTLQQQHQRLAQLLLILSHIRH 331
           |: ::||:|:: |||  ||| ||::|||:||
Qy     384 LVWVIAKSGISSQQQSMRLANLLMLLSHVRH 414


RESULT     2
ID   5223606-6      STANDARD;      PRT;    393 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Patent No. 5223606.
XX
CC   Patent No. 5223606
CC        APPLICANT:  BLAUDIN DE THE, HUGHES;MARCHIO, AGNES;TIOLLAIS,
CC   PIERRE;DEJEAN, ANNE
CC        TITLE OF INVENTION: STEROID/THYROID HORMONE RECEPTOR-RELATED
CC   PROTEIN INAPPROPRIATELY EXPRESSED IN HUMAN HEPATOCELLULAR CARCINOMA
CC        NUMBER OF SEQUENCES: 11
CC        CURRENT APPLICATION DATA:
CC          APPLICATION NUMBER:  US/07/134,130
CC          FILING DATE: 17-DEC-1987
CC        PRIOR APPLICATION DATA:
CC   SEQ ID NO:6:
CC          LENGTH: 363
SQ   SEQUENCE   393 AA; 44638 MW; 907537 CN;

  Query Match            21.2%;  Score 660;  DB 1;  Length 363;
  Best Local Similarity  32.2%;  Pred. No. 5.64e-50;
  Matches  104;  Conservative   88;  Mismatches 120;  Indels  11;  Gaps  11;

Db       1 CLICGDEASGCHYGVLTCGSCKVFFKRAMEGQHNYLCAGRNDCIVDKIRRKNCPACRLRK 60
           | :|:| ||| |||| :| :|| ||||:::|:::|:|:: |:| :|| |||:| ||||||
Qy      96 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRLRK 155

Db      61 CCQAGMVL-GGRKFK-KFNKVRVMRALDAVALPAPVGIPNESQRIIFSPSQEIQLIP-PL 117
           | : |||  |:|: :  :  ||  |:|       |  :  : ::           | |  |
Qy     156 CYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 215

Db     118 INLLMSIEPDVIYAGHDNTKPDTSSSLLTSLNQLGERQLLSVVKWSKSLPGFRNLHIDDQ 177
           :  |:  ||  :  :: :: | | :|:: ||  |....|. .. |.| .||| .| . ||
Qy     216 VLTLLEAEPPHVLISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFDQ 274

Db     178 ITLIQYSWMSLMVFGLGWRSYKHVSGQMLYFAPDLILNEQRMK-ESSFYSLCLTMWQIPQ 236
           : |::  || ::: || |||   | :| :: |||||:|: :   |     ::  :   :
Qy     275 VRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDMLLATTS 332

Db     237 EFVKLQVSQEEFLCMKVLLLLNT-I-PLEGLRSQSQFEEMRSSYIRELIKAIG-LRQKGV 293
            |  | : : |:||:| ::|||: : ||      :::    : :  :  :  |:  : |:
Qy     333 RFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKSG 392

Db     294 VSSSQRFYQLTKLLDNLHDLVKQ 316
           :|| |:  :|::||  |   |::
Qy     393 ISSQQQSMRLANLL-MLLSHVRH 414


RESULT     3
ID   5223606-7      STANDARD;      PRT;    386 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Patent No. 5223606.
XX
CC   Patent No. 5223606
CC        APPLICANT:  BLAUDIN DE THE, HUGHES;MARCHIO, AGNES;TIOLLAIS,
CC   PIERRE;DEJEAN, ANNE
CC        TITLE OF INVENTION: STEROID/THYROID HORMONE RECEPTOR-RELATED
CC   PROTEIN INAPPROPRIATELY EXPRESSED IN HUMAN HEPATOCELLULAR CARCINOMA
CC        NUMBER OF SEQUENCES: 11
CC        CURRENT APPLICATION DATA:
CC          APPLICATION NUMBER:  US/07/134,130
CC          FILING DATE: 17-DEC-1987
CC        PRIOR APPLICATION DATA:
CC   SEQ ID NO:7:
CC          LENGTH: 356
SQ   SEQUENCE   386 AA; 44008 MW; 864027 CN;

  Query Match            19.6%;  Score 609;  DB 1;  Length 356;
  Best Local Similarity  36.0%;  Pred. No. 3.36e-45;
  Matches   95;  Conservative   70;  Mismatches  84;  Indels  15;  Gaps  14;

Db       1 CLVCSDEASGCHYGVLTCGSCKVFFKRAVEGQHNYLCAGRNDCIIDKIRRKNCPACRYRK 60
           | |||| ||| |||| :| :|| ||||:::|:::|:|:: |:| ||| |||:| ||| ||
Qy      96 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRLRK 155

Db      61 CLQAGMNL-EARKTK-K-KIKGIQQATTG-V-SQETSENPGNKT-IVPATLPQ-LTPT-L 112
           | : ||    :|: :   ::   |:::    :        :  :|  :  |   | : |:|  |
Qy     156 CYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 215

Db     113 V-SLLEVIEPEVLYAGYDSSVPDSTWRIMTTLNMLGGRQVIAAVKWAKAIPGFRNLHLDD 171
           | :|||   | || :   ::  :::  :| :| :|: ::::  : ||| |||| :| | |
Qy     216 VLTLLEAEPPHVLISRPSAPFTEAS--MMMSLTKLADKELVHMISWAKKIPGFVELSLFD 273

Db     172 QNTLLQYSWM-FLMAALGWRSYRQSSANLLCFAPDLIIN-EQRMTLPCMYDQCKHMLYVS 229
           |  ||:  ||  || :| |||   ::: :|: ||||:::: :: : :  : :     :|  :
Qy     274 QVRLLESCWMEVLMMGLMWRSIDHPG-KLI-FAPDLVLDRDEGKCVEGILEIFDMLLATT 331

Db     230 SELHRLQVSYEEYLCMKYLLLLSS 253
           | :: | :    ||||:| ::||:|
Qy     332 SRFRELKLQHKEYLCVKAMILLNS 355


RESULT     4
ID   US-07-716-827C-5      STANDARD;      PRT;    795 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 5, Application US/07716827C.
XX
CC    Sequence 5, Application US/07716827C
CC    Patent No. 5215916
CC     GENERAL INFORMATION:
CC       APPLICANT:  Simons Jr., Stoney S.
CC       APPLICANT:  Yamamoto, K. R.
CC       APPLICANT:  Chakraborti, P. K.
CC       APPLICANT:  Garabedian, M. J.
CC       TITLE OF INVENTION:  SUPER GLUCOCORTICOID RECEPTORS
CC       NUMBER OF SEQUENCES:  5
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Cushman, Darby & Cushman
CC         STREET:  Eleventh floor, 1615 L Street, N.W.
CC         CITY:  Washington
CC         STATE:  D.C.
CC         COUNTRY:  U.S.A.
CC         ZIP:  20036-5601
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/716,827C
CC         FILING DATE:  19910619
CC         CLASSIFICATION:  435
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Scott, Watson T.
CC         REGISTRATION NUMBER:  26,581
CC         REFERENCE/DOCKET NUMBER:  WTS/5683/84453
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (202)861-3000
CC         TELEFAX:  (202)822-8944
CC         TELEX:  6714627 CUSH
CC     INFORMATION FOR SEQ ID NO:  5:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  795 amino acids
CC         TYPE:  AMINO ACID
CC         TOPOLOGY:  linear
SQ   SEQUENCE   795 AA; 87498 MW; 3326890 CN;

  Query Match            18.6%;  Score 577;  DB 3;  Length 795;
  Best Local Similarity  32.9%;  Pred. No. 3.25e-42;
  Matches   92;  Conservative   83;  Mismatches  89;  Indels  16;  Gaps  15;

Db      424 SPPSSSAATGPPPKIPTLVISDEASGCHYGVLTCGSCKVFFKRAVEGQHNYLCAGRNDCI 483
            ::| :::::       : .| |||| ||| |||| .| .|| ||||...|....|.|.. |.|
Qy       81 ASPVIGPGSK-RDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCT 139

Db      484 IDKIRRKNCPACRYRKCLQAGMNL-EARKTK-K-KIKGIQQATAG-V-SQDTSENPN-KT 537
            ||| |||:| ||| ||| : ||       .|.  .   ..   |...  .    .   .  .
Qy      140 IDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHA 199

Db      538 -IVPAALPQ-LTPT-LV-SLLEVIEPEVLYAGYDSSVPDSAWRIMTTLNMLGGRQVIAAV 593
             |   | : |:| || .|||  | ||  .   ..          : :| |
Qy      200 PRVRELLLDALSPEQLVLTLLEAEPPHVLISRPSAPFTEAS--MMMSLTKLADKELVHMI 257

Db      594 KWAKAILGLRNLHLDDQMTLLQYSWMFLMAFALGWRSYRQSSGNLLCFAPDLIIN-EQRM 652
             ||| | |: :| | ||. ||.  ||  .   :| |||   : : :  ||||    .
Qy      258 SWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHP-GKLI-FAPDLVLDRDEGK 315

Db      653 SLPCMYDQCKHMLFVSSELQRLQVSYEEYLCMKTLLLLSS 692
             :  :  :     :|  :| :: | :    ||||:|:::||:|
Qy      316 CVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNS 355

RESULT    5
ID   US-07-952-800 2     STANDARD;     PRT;   533 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 2, Application US/07952800.
XX
CC    Sequence 2, Application US/07952800
CC    Patent No. 5403925
CC     GENERAL INFORMATION:
CC       APPLICANT:  OZATO, KEIKO
CC       TITLE OF INVENTION:  A NEW MEMBER OF THE NUCLEAR HORMONE
CC       TITLE OF INVENTION:  RECEPTOR SUPERFAMILY AND A CDNA CLONE THEREOF
CC       NUMBER OF SEQUENCES:  4
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  TOWNSEND AND TOWNSEND
CC         STREET:  1 MARKET PLAZA, STEUART TOWER, SUITE 2000
CC         CITY:  SAN FRANCISCO
CC         STATE:  CA
CC         COUNTRY:  USA
CC         ZIP:  94105
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/952,800
CC         FILING DATE:  19920928
CC         CLASSIFICATION:  435
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Weber, Kenneth A
CC         REGISTRATION NUMBER:  31,677
CC         REFERENCE/DOCKET NUMBER:  15280-21-1
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  415-543-9600
CC         TELEFAX:  415-543-5043
```

```
CC     INFORMATION FOR SEQ ID NO:  2:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  533 amino acids
CC         TYPE:  AMINO ACID
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   533 AA; 56921 MW; 1406341 CN;

  Query Match            16.3%;  Score 506;  DB 5;  Length 533;
  Best Local Similarity  39.2%;  Pred. No. 1.28e-35;
  Matches  102;  Conservative   63;  Mismatches  98;  Indels  13;  Gaps  13;

Db      186 VRGLHCPPPPGGPGAGK--RLCAICGDRSSGKHYGVYSCEGCKGFFKRTIRKDLTYSCRD 243
            | | .|..|  |||. .   ..||:|:| :|| |||| |||||||:||||:|:      | |
Qy       75 VSGNRCASPVIGPGSKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPA 134

Db      244 NKDCIVDKRQRNRCQYCRYQKCLATGM-K----PE--AVQ-EEPQRGKDKDGDGEG-AGG 294
              ..||.||  .|. || || .||    ||  |    ||  . .   |||. | .    |
Qy      135 TNQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKR 194

Db      295 APEEMPVDR-ILRAELAVEQKS-DQ-GVEGPGGTGGSGSSP-NDP--VTNICQAADKQLF 348
            :    | | :|   |. ||           | |     .  |.|  ..  . ::   ||||:|
Qy      195 SGGHAPRVRELLLDALSPEQLVLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELV 254

Db      349 TLVFWAKRIPHFSSLPLDDQVILLRAGWNELLIASFSHRSIDVRDGILLATGLHVHRNSA 408
              .. ||||.|| |  |.| |||| ||  . | |.|. ..  ||||    .::| | : : :
Qy      255 HMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEG 314

Db      409 HSA-GVGAIFDRVLTELVSKMRDMRMDKTELGCLRAIILFNP 449
                |.  ||| .:.    |. |.....  |  |.:|:||:|:
Qy      315 KCVEGILEIFDMLLAT-TSRFRELKLQHKEYLCVKAMILLNS 355

RESULT    6
ID   US-07-952-800-4     STANDARD;     PRT;   446 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 4, Application US/07952800.
XX
CC    Sequence 4, Application US/07952800
CC    Patent No. 5403925
CC     GENERAL INFORMATION:
CC       APPLICANT:  OZATO, KEIKO
CC       TITLE OF INVENTION:  A NEW MEMBER OF THE NUCLEAR HORMONE
CC       TITLE OF INVENTION:  RECEPTOR SUPERFAMILY AND A CDNA CLONE THEREOF
CC       NUMBER OF SEQUENCES:  4
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  TOWNSEND AND TOWNSEND
CC         STREET:  1 MARKET PLAZA, STEUART TOWER, SUITE 2000
CC         CITY:  SAN FRANCISCO
CC         STATE:  CA
CC         COUNTRY:  USA
CC         ZIP:  94105
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/952,800
CC         FILING DATE:  19920928
CC         CLASSIFICATION:  435
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Weber, Kenneth A
CC         REGISTRATION NUMBER:  31,677
CC         REFERENCE/DOCKET NUMBER:  15280-21-1
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  415-543-9600
CC         TELEFAX:  415-543-5043
CC     INFORMATION FOR SEQ ID NO:  4:
```

```
CC        SEQUENCE CHARACTERISTICS:
CC          LENGTH: 446 amino acids
CC          TYPE: AMINO ACID
CC          TOPOLOGY: linear
CC        MOLECULE TYPE: protein
SQ    SEQUENCE   446 AA; 48023 MW; 964846 CN;

  Query Match             16.2%;  Score 505;  DB 5;  Length 446;
  Best Local Similarity   36.2%;  Pred. No. 1.58e-35;
  Matches  102;  Conservative   63;  Mismatches   98;  Indels   19;  Gaps   13;

Db       93 VRGLHCPPPPGGPGAGK--RLCAICGDRSSGKHYGVYSCEGCKGFFKRTIRKDLTYSCRD 156
            | | :|::|   |||: :   ::||:|:|  :||  ||||  |||||:|||||:|:      | |
Qy       75 VSGNRCASPVTGPGSKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPA 134

Db      157 NKDCTVDKRQRNRCQYCRYQKCLAIGM-K----RE--AVQ-EERQRGKDKDGDGD-AGG 207
             ::||:|| :|: || || :||    || |      ||  : :    |||: | :     | |
Qy      135 TNQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKR 194

        208 APEEMPVDR-ILEAELAVEQKS-DQ-GVEGPGATGGGGSSP-NDP--VTNICQAADKQLF 261
            :     |  |  :|    |: ||           | |    :  |:|  ::   : :'   |||:|
        195 SGGHAPRVRELLLDALSPEQLVLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELV 254

Db      262 TLVEWAKRIPHFSSLPLDDQVILLRAGWNELLIASFSHRSIDVRDGILLATGLHVHRNSA 321
             :: |||:|| |   |:| ||| || : | |:|: ::  ||||     :::|  | : |: .
Qy      255 HMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEG 314

Db      322 HSA-GVCAIFDRVLTELVSKMRDMRMDKTELGCLRAIILFNP 362
                |:  ||| :|:    |: |:::::  |   |::|:||:|:
Qy      315 KCVEGILEIFDMLLAT-TSRFRELKLQHKEYLCVKAMILLNS 355

RESULT     7
ID   PCT-US91 00399-2     STANDARD;     PRT;   462 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 2, Application PC/IUS9100399.
XX
CC    Sequence 2, Application PC/IUS9100399
CC     GENERAL INFORMATION:
CC        APPLICANT:  Mangelsdorf, Dr., David J.
CC        APPLICANT:  Evans Dr., Ronald M.
CC        TITLE OF INVENTION:  RETINOID RECEPTOR COMPOSITIONS AND
CC        TITLE OF INVENTION:  METHODS
          NUMBER OF SEQUENCES:  7
          CORRESPONDENCE ADDRESS:
            ADDRESSEE:  FITCH, EVEN, TABIN & FLANNERY
CC          STREET:  135 South LaSalle Street, Suite 900
CC          CITY:  Chicago
CC          STATE:  Illinois
CC          COUNTRY:  USA
CC          ZIP:  60603
CC        COMPUTER READABLE FORM:
CC          MEDIUM TYPE:  Floppy disk
CC          COMPUTER:  IBM PC compatible
CC          OPERATING SYSTEM:  PC-DOS/MS-DOS
CC          SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC        CURRENT APPLICATION DATA:
CC          APPLICATION NUMBER:  PCT/US91/00399
CC          FILING DATE:  19910122
CC          CLASSIFICATION:  435
CC        PRIOR APPLICATION DATA:
CC          APPLICATION NUMBER:  07/478,071
CC          FILING DATE:  09-FEB-1990
CC        ATTORNEY/AGENT INFORMATION:
CC          NAME:  Reiter, Mr., Stephen E.
CC          REGISTRATION NUMBER:  31192
CC          REFERENCE/DOCKET NUMBER:  50852
CC        TELECOMMUNICATION INFORMATION:
CC          TELEPHONE:  (619) 552-1311
```

```
CC          TELEFAX:  (619) 552-0095
CC          TELEX:  20 6566 PATLAW CGO
CC      INFORMATION FOR SEQ ID NO:  2:
CC        SEQUENCE CHARACTERISTICS:
CC          LENGTH:  462 amino acids
CC          TYPE:  AMINO ACID
CC          TOPOLOGY:  linear
CC        MOLECULE TYPE:  protein
SQ    SEQUENCE   462 AA; 50811 MW; 1063517 CN;

  Query Match             15.9%;  Score 493;  DB 9;  Length 462;
  Best Local Similarity   34.1%;  Pred. No. 2.03e 34;
  Matches   95;  Conservative   71;  Mismatches   94;  Indels   19;  Gaps   15;

Db      133 HICAICGDRSSGKHYGVYSCEGCKGFFKRTVRKDLTYTCRDNKDCLIDKRQRNRCQYCRY 192
            |:||:|:| :|| |||| |||||:||||:::      | |    ::| ||| :|: || ||
Qy       94 HFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRL 153

Db      193 QKCLAMGMKREAVQEER-Q-R-GK-DRNENE-VESTSSAN---EDMPVER-ILEAELAVE 243
            .||  .||  .  .  .||     |   .  .|. .|  .    ..|:         |  | :|   |: |
Qy      154 RKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPE 213

Db      244 PKTETYVEANMG--L-N-PSSP-NDP--VTNICQAADKQLFTLVEWAKRIPHFSELPLDD 296
                |  :|||:     |  :  ||:|  ::   :  ::    |||:|   :: |||:|| | ||:| |
Qy      214 QLVLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 273

Db      297 QVILLRAGWNELLIASFSHRSIAVKDGILLATGLHVHRNSAHSA-GVGAIFDRVLTELVS 355
            || || . | |.|. ..  |||       ...|  |  . |: :     |:  ||| :|:     |
Qy      274 QVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLAT-TS 332

Db      356 KMRDMQMDKTELGCLRAIVLFNPDSKGLSNPA-EVEALR 393
            : |:: ::   |   |::|::|:|:      |  :: : :: |
Qy      333 RFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSR 371

RESULT     8
ID   PCT-US91-00399-4     STANDARD;     PRT;   467 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 4, Application PC/IUS9100399.
XX
CC    Sequence 4, Application PC/IUS9100399
CC     GENERAL INFORMATION:
CC        APPLICANT:  Mangelsdorf, Dr., David J.
CC        APPLICANT:  Evans Dr., Ronald M.
CC        TITLE OF INVENTION:  RETINOID RECEPTOR COMPOSITIONS AND
CC        TITLE OF INVENTION:  METHODS
CC        NUMBER OF SEQUENCES:  7
CC        CORRESPONDENCE ADDRESS:
CC          ADDRESSEE:  FITCH, EVEN, TABIN & FLANNERY
CC          STREET:  135 South LaSalle Street, Suite 900
CC          CITY:  Chicago
CC          STATE:  Illinois
CC          COUNTRY:  USA
CC          ZIP:  60603
CC        COMPUTER READABLE FORM:
CC          MEDIUM TYPE:  Floppy disk
CC          COMPUTER:  IBM PC compatible
CC          OPERATING SYSTEM:  PC-DOS/MS-DOS
CC          SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC        CURRENT APPLICATION DATA:
CC          APPLICATION NUMBER:  PCT/US91/00399
CC          FILING DATE:  19910122
CC          CLASSIFICATION:  435
CC        PRIOR APPLICATION DATA:
CC          APPLICATION NUMBER:  07/478,071
CC          FILING DATE:  09-FEB-1990
CC        ATTORNEY/AGENT INFORMATION:
CC          NAME:  Reiter, Mr., Stephen E.
CC          REGISTRATION NUMBER:  31192
CC         REFERENCE/DOCKET NUMBER:  50852
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (619) 552-1311
CC         TELEFAX:  (619) 552-0095
CC         TELEX:  20 6566 PATLAW CGO
CC     INFORMATION FOR SEQ ID NO:  4:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  467 amino acids
CC         TYPE:  AMINO ACID
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   467 AA;  51216 MW;  1083042 CN;

  Query Match              15.9%;  Score 494;  DB 9;  Length 467;
  Best Local Similarity  34.1%;  Pred. No. 1.64e-34;
  Matches    95;  Conservative   71;  Mismatches  94;  Indels  19;  Gaps  15;

Db     138 HICAICGDRSSGKHYGVYSCEGCKGFFKRIVRKDLIYTCRDNKDCLIDKRQRNRCQYCRY 197
           |:||:|:| :|| |||  ||||||:||||:::       | |    ..| |||  |  || ||
Qy      94 HFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRL 153

Db     198 QKCLAMGMKREAVQEER-Q-R-GK-DRNENE-VESTSSAN---EDMP-VEKILEAELAVE 248
           :||  :|| : : : ||    |     |    |       |      | |   |
Qy     154 RKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPE 213

Db     249 PKTETYVEANMG-I-N-PSSP-NDP--VTNICQAADKQLFTLVEWAKRIPHFSELPLDD 301
              | :||:    | : ||:|  ::  : ::   |:|:|  ..  |||:||  | || | |
Qy     214 QLVLTLLEAEPPHVLISRRSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 273

Db     302 QVILLRAGWNELLIASFSHRSIAVKDGILLATGLHVHRNSAHSA-GVGAIFDRVLTELVS 360
           || || : | |:|: ::  |||         |   |      |  |||  ||    |
Qy     274 QVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLAT-TS 332

Db     361 KMRDMQMDKTELGCLRAIVLFNPDSKGLSNPA-EVEALR 398
           : |:: ::  |  |::|::|:|:      |  . . .. |
Qy     333 RFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSR 371

RESULT     9
ID   PCT-US91-00399-6      STANDARD;      PRT;    463 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 6, Application PC/IUS9100399.
XX
CC    Sequence 6, Application PC/IUS9100399
CC     GENERAL INFORMATION:
CC       APPLICANT:  Mangelsdorf, Dr., David J.
CC       APPLICANT:  Evans Dr., Ronald M.
CC       TITLE OF INVENTION:  RETINOID RECEPTOR COMPOSITIONS AND
CC       TITLE OF INVENTION:  METHODS
CC       NUMBER OF SEQUENCES:  7
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  FITCH, EVEN, TABIN & FLANNERY
CC         STREET:  135 South LaSalle Street, Suite 900
CC         CITY:  Chicago
CC         STATE:  Illinois
CC         COUNTRY:  USA
CC         ZIP:  60603
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US91/00399
CC         FILING DATE:  19910122
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/478,071
CC         FILING DATE:  09-FEB-1990
```

```
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Reiter, Mr., Stephen E.
CC         REGISTRATION NUMBER:  31192
CC         REFERENCE/DOCKET NUMBER:  50852
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (619) 552-1311
CC         TELEFAX:  (619) 552-0095
CC         TELEX:  20 6566 PATLAW CGO
CC     INFORMATION FOR SEQ ID NO:  6:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  463 amino acids
CC         TYPE:  AMINO ACID
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   463 AA;  50978 MW;  1097932 CN;

  Query Match              14.8%;  Score 461;  DB 9;  Length 463;
  Best Local Similarity  33.4%;  Pred. No. 1.81e-31;
  Matches    97;  Conservative   74;  Mismatches  96;  Indels  23;  Gaps  18;

Db     127 PSTSPGSLVK-HICAICGDRSSGKHYGVYSCEGCKGFFKRTIRKDLIYTCRDNKDCLIDK 185
           | |.|||    |:|||.|.| .|| |||| ||||||.||||.|.      | |    ::| |||
Qy      83 PVTSPGSKPLAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDK 142

Db     186 RQRNRCQYCRYQKCLVMGMKREAVQEER-Q-R--SRER-AESEAECASSS-HED--MP-V 236
            .|. || || .|| .||  . . . ||   |   |.| |: : ||: : :     | |
Qy     143 NRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRV 202

Db     237 -ERILEA---E-LAVEP-KTES-YGDMNVENS--TNDPVT-NICHAADKQLFTLVEWAKR 286
            | .|.|   | |  .    .|.    ..  ..  |: ::  ::   |||:|  :: |||:
Qy     203 RELLLDALSPEQLVLTLLEAEPPHVLISRRSAPFTEASMMMSLIKLADKELVHMISWAKK 262

Db     287 IPHFSDLTLEDQVILLRAGWNELLIASFSHRSVSVQDGILLAIGLHVHRSSAHS-RGVGS 345
           || | .|:| ||| || : | |:|: ::  ||:       :::|  | : |  :     ::
Qy     263 IPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILE 322

Db     346 IFDRVLTELVSKMKDMQMDKSELGCLRAIVLFNPDAKGLSNPS-EVETLR 394
           ||| :|:    |: ::: ::  |  |::|::|:|:     |  :: : :: |
Qy     323 IFDMLLAT-TSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSR 371

RESULT    10
ID   US-08-333-358-8      STANDARD;      PRT;    440 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 8, Application US/08333358.
XX
CC    Sequence 8, Application US/08333358
CC    Patent No. 5571696
CC     GENERAL INFORMATION:
CC       APPLICANT:  EVANS Ph.D., RONALD M.
CC       APPLICANT:  MANGELSDORF Ph.D., DAVID J.
CC       APPLICANT:  ONG Ms., ESTELITA S.
CC       APPLICANT:  ORO Ph.D., ANTHONY E.
CC       APPLICANT:  BORGMEYER Ph.D., UWE K.
CC       APPLICANT:  GIGUERE Ph.D., VINCENT NMN
CC       APPLICANT:  YAO Mr., TSO-PANG NMN
CC       TITLE OF INVENTION:  NOVEL RECEPTORS
CC       NUMBER OF SEQUENCES:  14
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Pretty, Schroeder, Brueggemann & Clark
CC         STREET:  444 So. Flower St., Suite 2000
CC         CITY:  Los Angeles
CC         STATE:  CA
CC         COUNTRY:  US
CC         ZIP:  90071-2921
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC          SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC        CURRENT APPLICATION DATA:
CC          APPLICATION NUMBER: 08/779,333,359
CC          FILING DATE:
CC          CLASSIFICATION:  435
CC        PRIOR APPLICATION DATA:
CC          APPLICATION NUMBER:  US/07/761,068
CC          FILING DATE:
CC        ATTORNEY/AGENT INFORMATION
CC          NAME:  Reiter Ph.D., Stephen E.
CC          REGISTRATION NUMBER:  31192
CC          REFERENCE/DOCKET NUMBER:  P31 8936
CC        TELECOMMUNICATION INFORMATION:
CC          TELEPHONE:  (619) 535-9001
CC          TELEFAX:  (619) 535-8949
CC      INFORMATION FOR SEQ ID NO:  8:
CC        SEQUENCE CHARACTERISTICS:
CC          LENGTH:  440 amino acids
CC          TYPE:  amino acid
            TOPOLOGY:  linear
        MOLECULE TYPE:  protein
     SEQUENCE   440 AA; 49883 MW; 957063 CN;

  Query Match             13.5%;  Score 420;  DB 6;  Length 440;
  Best Local Similarity  30.8%;  Pred. No. 1.05e-27;
  Matches    84;  Conservative   68;  Mismatches 104;  Indels  17;  Gaps  15;

Db      80 GPAPKMLGNELCSVCGDKASGFHYNVLSCEGCKGFFRRSVIKGAHYICHSGGHCPMDTYM 139
           ||::|     :|:||:| |||:|| | ||||||:||:||:        ||| :   :| :|
Qy      86 GPGSKR-DAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNR 144

Db     140 RRKCQFCRLRKCRQAGMREECVLSEEQ--IRLKKLKRQEEEQAHATSLPPRRSSPPQILP 197
           |: || |||||| : ||     |      |:    || :   |   :|| | :: : | :: :   :
Qy     145 RKSCQACRLRKCYEVGM-VKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVR 203

Db     198 QLSEEQICMIEKIVAAQQQCNERRSESFEIRVTPWPMAHEPHSFEARQQRFAHFTELAIVS 257
           :|   : |:   | || :  :  :   :  :  | :: :| :: | : ||      : :| ||
Qy     204 ELLLDALS-PEQLVLILLEAE-P-PHVL-IS-RP-SA-FT-EASM-MMS-LTKLADKE 252

Db     258 VQEIVDFAKQLPGFLQLSREDQIALLKTSAIEVMLLETSRRYNPGSESITFLKDFSYNRE 317
           :  ::   || :|||::||    ||: || :   :||:::     :|       :  :  |   |:  :|:
Qy     253 LVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRD 312

Db     318 DFAKAGLQV-EFINPIFEFSRAMNELQLNDAEF 349
           : :|     : |::: ::  :     || |    |:
Qy     313 E-GKCVEGILEIFDMLLATTSRFRELKLQHKEY 344

   ULT   11
     5260432-2      STANDARD;      PRT;    492 AA.
XX
AC    xxxxxx
XX
DT    01-JAN-1900
XX
DE    Patent No. 5260432.
XX
CC    Patent No. 5260432
CC         APPLICANT: TAKAKU, FUMIMARO;ISHIKAWA, TAKASHI;IMAWARI, MICHIO;
CC    EVANS, RONALD M.;UMESONO, KAZUHIKO
CC         TITLE OF INVENTION: HUMAN GAMMA RETINOIC ACID RECEPTOR DNA
CC         NUMBER OF SEQUENCES: 4
CC         CURRENT APPLICATION DATA:
CC           APPLICATION NUMBER:  US/07/370,407
CC           FILING DATE: 22-JUN-1989
CC    SEQ ID NO:2:
CC           LENGTH: 454
SQ    SEQUENCE   492 AA; 54527 MW; 1333563 CN;

  Query Match             13.2%;  Score 411;  DB 1;  Length 454;
  Best Local Similarity  29.7%;  Pred. No. 6.95e-27;
  Matches    74;  Conservative   61;  Mismatches 102;  Indels  12;  Gaps  12;
```

```
Db      90 CFVCNDKSSGYHYGVSSCEGCKGFFRRSIQKNMVYTCHRDKNCIINKVTRNRCQYCRLQK 149
           | ||:| :||||||| ||||||:||:||||  :  | |     : | |:|   |: || |||:|
Qy      96 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 155

Db     150 CFEVGMSKEAVRNDR-N-K-KKKEVKEEGSPDSYELSPQLEELITKVSKAHQETF-PS-L 204
           |:||||| | : | :|    :   :::     :              : :       :|      ::: |  |
Qy     156 CYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 215

Db     205 C-QL-GKYTTNSSADHR-VQL-DLGLWDKFSELATKCIIKIVEFAKRLPGFTGLSIADQI 260
              |        :       :        :  :  ::     :: || | :: ::  ||::|||  :|: ||:
Qy     216 VLILLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFDQV 275

Db     261 TLLKAAC-LDILMLRICTRYTPEQDTMTFSDGLTLNRTQ-MHNAGFGPLTDLVFAFAGQL 318
            ||  :| ::::||: :  |           : |: | |:| : :   |:  : |:::: ::::
Qy     276 RLL-ESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSRF 334

Db     319 LPLEMDDTE 327
            | ::  |
Qy     335 RELKLQHKE 343


RESULT   12
ID   5223606-3      STANDARD;      PRT;    398 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Patent No. 5223606.
XX
CC   Patent No. 5223606
CC        APPLICANT:  BLAUDIN DE THE, HUGHES;MARCHIO, AGNES;TIOLLAIS,
CC   PIERRE;DEJEAN, ANNE
CC        TITLE OF INVENTION: STEROID/THYROID HORMONE RECEPTOR-RELATED
CC   PROTEIN INAPPROPRIATELY EXPRESSED IN HUMAN HEPATOCELLULAR CARCINOMA
CC        NUMBER OF SEQUENCES: 11
CC        CURRENT APPLICATION DATA:
CC          APPLICATION NUMBER  US/07/134,130
CC          FILING DATE: 17-DEC-1987
CC        PRIOR APPLICATION DATA:
CC   SEQ ID NO:3:
CC          LENGTH: 368
SQ   SEQUENCE   398 AA; 44842 MW; 907111 CN;

  Query Match             13.1%;  Score 407;  DB 1;  Length 368;
  Best Local Similarity  28.9%;  Pred. No. 1.61e-26;
  Matches    72;  Conservative   66;  Mismatches  99;  Indels  12;  Gaps  12;

Db       1 CFVCQDKSSGYHYGVSACEGCKGFFRRSIQKNMIYTCHRDKNCVINKVTRNRCQYCRLQK 60
           | || | :||||||| :|||||:||:|||| :   | |     : | |:|  |: || |||:|
Qy      96 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 155

Db      61 CFEVGMSKESVRNDR-N-K-KKKETSKQECTESYEMTAELDDLTEKIRKAHQETF-PS-L 115
           |:||||| | : | :|    :   ::: | :|           ::          : ::|      ::: |  |
Qy     156 CYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 215

Db     116 C-QL-GKYTTNSSADHR-VRL-DLGLWDKFSELATKCIIKIVEFAKRLPGFTGLTIADQI 171
              |        :       :        :  :  ::     :: || | :: ::  ||::|||  |:: ||:
Qy     216 VLILLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFDQV 275

Db     172 ILLKAAC-LDILILRICIRYTPEQDTMTFSDGLTLNRTQ-MHNAGFGPLTDLVFTFANQL 229
            ||  :| ::::|:: :  |           : |: | |:| : :   |:  : |:::: ::::
Qy     276 RLL-ESCWMEVLMMGLMWRSIDHPGKLIFAPDLVIDRDEGKCVEGILEIFDMLLATTSRF 334

Db     230 LPLEMDDTE 238
            | ::  |
Qy     335 RELKLQHKE 343


RESULT   13
ID   5223606-2      STANDARD;      PRT;    485 AA.
XX
AC   xxxxxx
```

```
XX
DT   01-JAN-1900
XX
DE   Patent No. 5223606
XX
CC   Patent No. 5223606
CC       APPLICANT.  BLAUDIN DE THE, HUGHES;MARCHIO, AGNES;TIOLLAIS,
CC   PIERRE;DEJEAN, ANNE
CC       TITLE OF INVENTION. STEROID/THYROID HORMONE RECEPTOR-RELATED
CC   PROTEIN INAPPROPRIATELY EXPRESSED IN HUMAN HEPATOCELLULAR CARCINOMA
CC       NUMBER OF SEQUENCES: 11
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER: [illegible]
CC         FILING DATE: 17-DEC-1987
CC       PRIOR APPLICATION DATA:
CC   SEQ ID NO:2:
CC         LENGTH: 448
SQ   SEQUENCE   485 AA; 54437 MW; 1345912 CN;

 uery Match             13.1%;  Score 407;  DB 1;  Length 448;
 est Local Similarity  28.9%;  Pred. No. 1.61e-26;
 atches    72;  Conservative   66;  Mismatches   99;  Indels  12;  Gaps  12;

Db     81 CFVCQDKSSGYHYGVSACEGCKGFFRRSIQKNMIYTCHRDKNCVINKVTRNRCQYCRLQK 140
           |  || |  :||||||||| :|||||.||:||||  :   | |     : | |:|  |: || |||:|
Qy     96 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 155

Db    141 CFEVGMSKESVRNDR N-K-KKKETSKQECTESYEMTAELDDLTEKIRKAHQETF-PS-L 195
           |:||||| | : | :|    :  ::: | :|         ::         : ::|      ::: |  |
Qy    156 CYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 215

Db    196 C-QI-GKYTTNSSADHR-VRL-DLGLWDKFSELATKCIIKIVEFAKRLPGFTGLTIADQI 251
              |       :     :      : : ::     :: || | . . . .   ||:.|||  |.. ||:
Qy    216 VLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFDQV 275

Db    252 TLLKAAC-LDILILRICTRYTPEQDTMTFSDGLTLNRTQ-MHNAGFGPLTDLVFTFANQL 309
            ||  :|  :::|::  :  |           . |.   | |.|  . .     |.  : |.... ....:
Qy    276 RLL-ESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSRF 334

Db    310 LPLEMDDTE 318
            | ::  |
Qy    335 RELKLQHKE 343


RESULT   14
ID   PCT-US92-02320A-4      STANDARD;      PRT;   462 AA.
XX
     xxxxxx

     01-JAN-1900
XX
DE   Sequence 4, Application PC/TUS9202320A.
XX
CC    Sequence 4, Application PC/TUS9202320A
CC     GENERAL INFORMATION:
CC       APPLICANT:  Sloan-Kettering Institute  For Cancer Research
CC       TITLE OF INVENTION:  METHODS FOR DETECTION AND TREATMENT OF CANCER
CC       NUMBER OF SEQUENCES:  6
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  COOPER & DUNHAM
CC         STREET:  30 ROCKEFELLER PLAZA
CC         CITY:  NEW YORK
CC         STATE:  NEW YORK
CC         COUNTRY:  U.S.A.
CC         ZIP:  10112
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.24
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US92 02320A
CC         FILING DATE:
CC         CLASSIFICATION:
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 675,084
CC         FILING DATE.  22-MAR 1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 673,838
CC         FILING DATE:  22-MAR-1991
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  WHITE, JOHN P
CC         REGISTRATION NUMBER:  28,678
CC         REFERENCE/DOCKET NUMBER:  38694-PCT
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE.  (212) 977-9550
CC         TELEFAX.  (212) 664 0525
CC         TELEX.  (212) 422523 COOP UI
CC     INFORMATION FOR SEQ ID NO:  4:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  462 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE.  protein
SQ   SEQUENCE   462 AA, 50771 MW, 1101159 CN;

  Query Match             12.6%,  Score 393,  DB 10;  Length 462;
  Best Local Similarity  29.7%;  Pred. No. 3.04e-25;
  Matches    74,  Conservative   60,  Mismatches 103,  Indels  12,  Gaps  12;

Db     89 CFVCQDKSSGYHYGVSACEGCKGFFRRSIQKNMVYTCHRDKNCIINKVTRNRCQYCRLQK 147
           |  || |  :||||||||| :|||||:||:||||  :   | |     : | |:|   |: || |||:|
Qy     96 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 155

Db    148 CFEVGMSKESVRNDR-N-K-KKKEVPKPECSESYTLTPEVGELIEKVRKAHQETF-PA-L 202
           |:|||| | : | :|    :  ::: :  |          :    |     :||     ::: |  |
Qy    156 CYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 215

Db    203 C-QL-GKYTTNN-SSEQRVSL-DIDLWDKFSELSTKCIIKTVEFAKQLPGFTTLTIADQI 258
              |      .    |       .. .  .    .. |. | .:   :  || :|||  |:: ||:
Qy    216 VLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFDQV 275

Db    259 TLLKAAC-LDILILRICTRYTPEQDTMTFSDGLTLNRTQ-MHNAGFGPLTDLVFAFANQL 316
            ||  :| ::.|.. :  |          . |.  | |.| : :   |:  : |:::| ::::
Qy    276 RLL-ESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSRF 334

Db    317 LPLEMDDAE 325
            | ::  |
Qy    335 RELKLQHKE 343


RESULT   15
ID   PCT-US92-02320A-2      STANDARD;      PRT;   797 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 2, Application PC/TUS9202320A.
XX
CC    Sequence 2, Application PC/TUS9202320A
CC     GENERAL INFORMATION.
CC       APPLICANT   Sloan-Kettering Institute, For Cancer Research
CC       TITLE OF INVENTION.  METHODS FOR DETECTION AND TREATMENT OF CANCER
CC       NUMBER OF SEQUENCES:  6
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  COOPER & DUNHAM
CC         STREET:  30 ROCKEFELLER PLAZA
CC         CITY:  NEW YORK
CC         STATE.  NEW YORK
CC         COUNTRY:  U.S.A.
CC         ZIP:  10112
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE.  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.24
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US92/02320A
CC         FILING DATE:
CC         CLASSIFICATION:
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 675,084
CC         FILING DATE:  22-MAR-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 673,838
CC         FILING DATE:  22-MAR-1991
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  WHITE, JOHN P
CC         REGISTRATION NUMBER:  28,678
CC         REFERENCE/DOCKET NUMBER:  38694-PCT
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 977-9550
CC         TELEFAX:  (212) 644-0525
CC         TELEX:  (212) 422523 COOP UI
      INFORMATION FOR SEQ ID NO:  2:
         SEQUENCE CHARACTERISTICS:
           LENGTH:  797 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   797 AA;  89281 MW;  3152972 CN;

  Query Match            12.6%;  Score 393;  DB 10;  Length 797;
  Best Local Similarity  29.7%;  Pred. No. 3.04e-25;
  Matches    74;  Conservative   60;  Mismatches 103;  Indels  12;  Gaps  12;

Db    423 CFVCQDKSSGYHYGVSACEGCKGFFRRSIQKNMVYICHRDKNCIINKVIRNRCQYCRLQK 482
          | || | :||||||| :|||||:||:||||  : | |    :| |:|  |: || |||:|
Qy     96 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCIIDKNRRKSCQACRLRK 155

Db    483 CFEVGMSKFSVRNDR-N-K-KKKEVPKPECSESYTLTPEVGELIEKVRKAHQETF-PA-L 537
          |:|||| | | :|   :  :·· ·  |       ·   |    ·||    ··· |  |
Qy    156 CYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 215

Db    538 C-QL-GKYTTNN-SSEQRVSL-DIDLWDKFSELSTKCIIKTVEFAKQLPGFTTLTIADQI 593
             |      :   |    :: :  :   :: |: | ::  :  || :|||  |:: ||:
Qy    216 VLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFDQV 275

Db    594 TLLKAAC-LDILILRICTRYTPEQDTMTFSDGLTLNRTQ-MHNAGFGPLTDLVFAFANQL 651
           ||  :| :::|:: :  |       : |:  | |·|       |     :   :   ··
Qy    276 RLL-ESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSRF 334

Db    652 LPLEMDDAE 660
            | ::  |
      335 RELKLQHKE 343

Search completed: Thu Dec 18 08:57:37 1997
Job time : 9 secs.
```

```
*************************************************************************
                         MPSRCH (TM)
*************************************************************************

      Release 2.1D John F. Collins, Biocomputing Research Unit
      Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
              Distribution rights by IntelliGenetics, Inc.

MPsrch_pp   protein - protein database search, using Smith-Waterman algorithm

 on:             Thu Dec 18 09:46:19 1997; MasPar time 5.56 Seconds
                                           414.836 Million cell updates/sec
Tabular output not generated.

Title:           >US-08-826-361A-5
Description:     (1-477) from US08826361A.pep
Perfect Score:   3532
Sequence:        1 MNYSIPSNVTNLEGGPGRQT..........FCSPAEDSKSKEGSQNPQSQ 477

Scoring table:   PAM 150
                 Gap 11

Searched:        54101 seqs, 4836506 residues

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database:        a-issued
                   1:backl 2:51 3:52 4:53 5:54 6:55 7:56 8:PCT90 9:PCT91
                   10:PCT92 11:PCT93 12:PCT94 13:PCT95 14:PCT96

Statistics:      Mean 32.905; Variance 149.394; scale 0.220

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES
                %
sult          Query
 No.  Score   Match Length DB  ID              Description              Pred. No.
---------------------------------------------------------------------------------
   1   1648    46.7    410  1  5223606-5       Patent No. 5223606.       8.78e-141
   2    660    18.7    362  1  5223606-4       Patent No. 5223606.       1.09e-48
   3    609    17.2    356  1  5223606-7       Patent No. 5223606.       5.02e-44
   4    577    16.3    795  3  US-07-716-      Sequence 5, Applicatio    4.15e-41
   5    533    15.1    533  5  US-07-952-      Sequence 2, Applicatio    4.13e-37
   6    529    15.0    446  5  US-07-952-      Sequence 4, Applicatio    9.51e-37
   7    514    14.6    467  9  PCT-US91-0      Sequence 4, Applicatio    2.17e-35
   8    512    14.5    462  9  PCT-US91-0      Sequence 2, Applicatio    2.57e-35
   9    486    13.8    463  9  PCT-US91-0      Sequence 6, Applicatio    7.35e-33
  10    441    12.5    454  1  5260432-2       Patent No. 5260432.       8.17e-29
  11    439    12.4    368  1  5223606-3       Patent No. 5223606.       1.23e-28
  12    439    12.4    449  1  5223606-2       Patent No. 5223606.       1.23e-28
  13    428    12.1    462 10  PCT-US92-0      Sequence 4, Applicatio    1.19e-27
  14    428    12.1    797 10  PCT-US92-0      Sequence 2, Applicatio    1.19e-27
  15    420    11.9    440  6  US-08-333-      Sequence 8, Applicatio    6.18e-27
  16    413    11.7    462  1  5171671-2       Patent No. 5171671.       2.61e-26
  17    361    10.2    469 13  PCT-US95-1      Sequence 2, Applicatio    1.08e-21
  18    357    10.1     89  6  US-08-561-      Sequence 3, Applicatio    2.43e-21
  19    357    10.1     89  5  US-08-087-      Sequence 3, Applicatio    2.43e-21
  20    352    10.0    484 13  PCT-US95-1      Sequence 1, Applicati     5.47e-21
  21    349     9.9     66  1  5217867-1       Patent No. 5217867.       1.23e-20
  22    329     9.3    451 13  PCT-US95-1      Sequence 2, Applicatio    7.02e-19
  23    315     8.9     70  3  US-07-737-      Sequence 3, Applicatio    1.18e-17
  24    310     8.8    461 13  PCT-US95-1      Sequence 2, Applicatio    3.21e-17
  25    310     8.8    461 13  PCT-US95-1      Sequence 2, Applicatio    3.21e-17
  26    310     8.8    461  7  US-08-330-      Sequence 2, Applicatio    3.21e-17
  27    310     8.8    461  7  US-08-330-      Sequence 2, Applicatio    3.21e-17
  28    307     8.7    446  7  US-08-342-      Sequence 4, Applicatio    5.85e-17
  29    307     8.7    446 13  PCT-US95-1      Sequence 3, Applicatio    5.85e-17
  30    305     8.6    544  6  US-08-333-      Sequence 2, Applicatio    9.92e-17
  31    305     8.6    410  1  5438126-2       Patent No. 5438126.       1.00e-16
  32    304     8.6    517  7  US-08-464       Sequence 2, Applicatio    1.07e-16
  33    304     8.6    558  6  US-08-333-      Sequence 12, Applicati    1.07e-16
  34    299     8.5    460  7  US-08-342-      Sequence 2, Applicatio    2.90e-16
  35    300     8.5    601  6  US-08-333-      Sequence 14, Applicati    2.37e-16
  36    297     8.4    355  1  5223606-4       Patent No. 5223606.       4.32e-16
  37    282     8.0    437 12  PCT-US94-0      Sequence 2, Applicatio    8.57e-15
  38    265     7.5    439  6  US-08-333-      Sequence 10, Applicati    2.48e-13
  39    238     6.7     89  5  US-08-087-      Sequence 2, Applicatio    4.93e-11
  40    238     6.7     89  6  US-08-561-      Sequence 2, Applicatio    4.93e-11
  41    191     5.4     60 12  PCT-US94-1      Sequence 1, Applicatio    4.05e-07
  42    169     4.8    367  3  US-07-737-      Sequence 4, Applicatio    1.46e-05
  43    141     4.0    272 13  PCT-US95-0      Sequence 23, Applicati    3.94e-03
  44    141     4.0    272  6  US-08-485-      Sequence 23, Applicati    3.94e-03
  45     97     2.7    298  6  US-08-099-      Sequence 4, Applicatio    6.68e+00

                              ALIGNMENTS

RESULT     1
ID   5223606-5     STANDARD;     PRT;   444 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Patent No. 5223606.
XX
CC   Patent No. 5223606
CC       APPLICANT: BLAUDIN DE THE, HUGHES;MARCHIO, AGNES;TIOLLAIS,
CC   PIERRE,DEJEAN, ANNE
CC       TITLE OF INVENTION: STEROID/THYROID HORMONE RECEPTOR-RELATED
CC   PROTEIN INAPPROPRIATELY EXPRESSED IN HUMAN HEPATOCELLULAR CARCINOMA
CC       NUMBER OF SEQUENCES: 11
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER: US/07/134,130
CC         FILING DATE: 17-DEC-1987
CC       PRIOR APPLICATION DATA:
CC   SEQ ID NO:5:
CC         LENGTH: 410
SQ   SEQUENCE   444 AA; 50020 MW; 1110634 CN;

  Query Match              46.7%;  Score 1648;  DB 1,  Length 410;
  Best Local Similarity   56.9%;  Pred. No. 8.78e-141;
  Matches  224;  Conservative   32;  Mismatches  64;  Indels  14;  Gaps  11;

Db        1 CAVCNDYASGYHYGVWSCEGCKAFFKRSIQGHNDYMCPATNQCTIDKNRRKSCQACRLRK 60
            ||||:||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Qy       96 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 155

Db       61 CYEVGMMKGGIRKDRRGGRMLKHKRQRDDGEGRGEVGSAGDMRAANLWPSPLMIKRSKKN 120
            ||||||:| | |::| | |::: |||   | :   :: : :|::     :|   : |
Qy      156 CYEVGMVKCGSRRERCGYRLVR--RQRSADE-QLHCAGKA-KRSGG--HAP-RV-RELL- 206

Db      121 SLALSLTADKMVSALLDAEPPILYSEYDPTRPFSEASMMGLLTNLADRELVHMINWAKRVP 180
              |  .:|...|:| :||:|||         |: ||:|||||  ||:|||:||||||:|||:::
Qy      207 -LD-ALSPEQLVLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIP 264

Db      181 GFVDLTLHDQVHLLECAWLEILMIGLVWRSMEHPVKLLFAPNLLLDRNQGKCVEGMVEIF 240
            ||||:|.| |||.||| |.|.||.||:|||..|| ||.|||.|:|||::|||||||::|||
Qy      265 GFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIF 324

Db      241 DMLLATSSRFRMMNLQGEEFVCLKSIILLNSGVYTFLSSTLKSLEEKDHIHRVLDKITDT 300
            ||||||:||||  ::||  |::|:|::||||||::| ::::|   :  :     : ::|: :||:
Qy      325 DMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSRKLAHLLNAVTDA 383
```

```
Db        301 LIHLMAKAGLTLQQQHQRLAQLLLILSHIRHMSNKGMEHLYSMKCKNVVPLYDLLLEMLD 360
              |: ::||.|.. |||  ||| ||..|||.|| ||||||||  .||||||||| ||||||||
Qy        384 LVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLN 443

Db        361 AHRLHA-PTSRGGASVEETDQSHLATAGSTSSHS 393
              || |::  :|  |:     :::|  :  || :::|
Qy        444 AHVLRGCKSSITGSECSPAEDSK-SKEGSQNPQS 476


RESULT     2
ID   5223606-6      STANDARD;      PRT;    393 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Patent No. 5223606.
XX
CC   Patent No. 5223606
CC       APPLICANT:  BLAUDIN DE THE, HUGHES;MARCHIO, AGNES;TIOLLAIS,
CC   PIERRE;DEJEAN, ANNE
CC       TITLE OF INVENTION: STEROID/THYROID HORMONE RECEPTOR-RELATED
CC   PROTEIN INAPPROPRIATELY EXPRESSED IN HUMAN HEPATOCELLULAR CARCINOMA
CC       NUMBER OF SEQUENCES: 11
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/134,130
CC         FILING DATE: 17-DEC-1987
CC       PRIOR APPLICATION DATA:
CC   SEQ ID NO:6:
CC         LENGTH: 363
SQ   SEQUENCE   393 AA; 44638 MW; 907537 CN;

  Query Match             18.7%;  Score 660;  DB 1;  Length 363;
  Best Local Similarity   32.2%;  Pred. No. 1.09e-48;
  Matches   104;  Conservative   88;  Mismatches 120;  Indels   11;  Gaps   11;

Db          1 CLICGDFASGCHYGVLTCGSCKVFFKRAMEGQHNYLCAGRNDCIVDKIRRKNCPACRLRK 60
              | :|:| ||| |||| :| :|| ||||::·|···|·|·· |·| ·|| |||·| ||||||
Qy         96 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 155

Db         61 CCQAGMVL-GGRKFK-KFNKVRVMRALDAVALPAPVGIPNESQRITFSPSQEIQLIP-PL 117
              | : |||  ':|: :  :  ||  |: |       |  :  : ::           | |  |
Qy        156 CYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 215

Db        118 INLLMSIEPDVIYAGHDNTKPDTSSSLLTSLNQLGERQLLSVVKWSKSLPGFRNLHIDDQ 177
              :  |:  ||  :  :: :: | | :|:: ||  |:···|· ·· |·| ·||| ·| · ||
Qy        216 VLTLLEAEPPHVLISRPSA-PFTFASMMMSLTKLADKELVHMISWAKKIPGFVELSLFDQ 274

Db        178 ITLIQYSWMSLMVFGLGWRSYKHVSGQMLYFAPDLILNEQRMK-ESSFYSLCLTMWQIPQ 236
              : |::  || ::: || |||  | :| :: |||||:|: :  |     ::  :   :
Qy        275 VRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDMLLATTS 332

Db        237 EFVKLQVSQEEFLCMKVLLLLNT-I-PLEGLRSQSQFEEMRSSYIRELIKAIG-LRQKGV 293
               |  |  :  : |:||:| ::|||: : ||      :::  :  :  :  :  |:  :  |:
Qy        333 RFRELKLQHKEYLCVKAMILLNSSMYPLVTAIQDADSSRKLAHLLNAVTDALVWVIAKSG 392

Db        294 VSSSQRFYQLTKLLDNLHDLVKQ 316
              :|| |:  :|::||  |   |::
Qy        393 ISSQQQSMRLANLL-MLLSHVRH 414


RESULT     3
ID   5223606-7      STANDARD;      PRT;    386 AA
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Patent No. 5223606.
XX
CC   Patent No. 5223606
CC       APPLICANT:  BLAUDIN DE THE, HUGHES;MARCHIO, AGNES;TIOLLAIS,
CC   PIERRE;DEJEAN, ANNE
CC       TITLE OF INVENTION: STEROID/THYROID HORMONE RECEPTOR-RELATED
CC   PROTEIN INAPPROPRIATELY EXPRESSED IN HUMAN HEPATOCELLULAR CARCINOMA
CC       NUMBER OF SEQUENCES: 11
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/134,130
CC         FILING DATE: 17-DEC-1987
CC       PRIOR APPLICATION DATA:
CC   SEQ ID NO:7:
CC         LENGTH: 356
SQ   SEQUENCE   386 AA; 44008 MW; 864027 CN;

  Query Match             17.2%;  Score 609;  DB 1;  Length 356;
  Best Local Similarity   36.0%;  Pred. No. 5.02e-44;
  Matches    95;  Conservative   70;  Mismatches   84;  Indels   15;  Gaps   14;

Db          1 CIVCSDEASGCHYGVLTCGSCKVFFKRAVEGQHNYLCAGRNDCIIDKIRRKNCPACRYRK 60
              | |||| ||| |||| .| .|| ||||..:|:::|:|:: |:| ||| |||:| ||| ||
Qy         96 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 155

Db         61 CLQAGMNL-EARKIK-K-KIKGIQQAIIG-V-SQEISENPGNKI-IVPAILPQ-LIPT-L 112
              | : ||    :|: :   ::   |:::    :        :  :|  :  |    | : |:|  |
Qy        156 CYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 215

Db        113 V-SLLEVIEPEVLYAGTDSSVPDSTWRIMTTLNMLGGRQVIAAVKWAKAIPGFRNLHLDD 171
              | ·|||   | ||  ·    ··   ···   ·| :| ·|· ::::  : ||| |||| :| | |
Qy        216 VLTLLEAEPPHVLISRPSAPFTEAS--MMMSLTKLADKELVHMISWAKKIPGFVELSLFD 273

Db        172 QNTLLQYSWM-FLMAALGWRSYRQSSANLLCFAPDLIIN-EQRMTLPCMYDQCKHMLYVS 229
              |  ||:  ||  || :| |||  ::: :|: |||||::: :: : :  : :   :|  :
Qy        274 QVRLLESCWMEVLMMGLMWRSIDHPG-KLI-FAPDLVLDRDEGKCVEGILEIFDMLLATT 331

Db        230 SELHRLQVSYEEYLCMKYLLLLSS 253
              | :: | :   ||||:| ::||:|
Qy        332 SEFRELKLQHKEYLCVKAMILLNS 355


RESULT     4
ID   US-07-716-827C-5      STANDARD;      PRT;    795 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 5, Application US/07716827C.
XX
CC   Sequence 5, Application US/07716827C
CC   Patent No. 5215916
CC    GENERAL INFORMATION:
CC       APPLICANT:  Simons Jr., Stoney S.
CC       APPLICANT:  Yamamoto, K. R.
CC       APPLICANT:  Chakraborti, P. K.
CC       APPLICANT:  Garabedian, M. J.
CC       TITLE OF INVENTION:  SUPER GLUCOCORTICOID RECEPTORS
CC       NUMBER OF SEQUENCES:  5
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Cushman, Darby & Cushman
CC         STREET:  Eleventh floor, 1615 L Street, N.W.
CC         CITY:  Washington
CC         STATE:  D.C.
CC         COUNTRY:  U.S.A.
CC         ZIP:  20036-5601
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/716,827C
CC         FILING DATE:  19910619
CC         CLASSIFICATION:  435
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Scott, Watson T.
CC         REGISTRATION NUMBER:  26,581
CC         REFERENCE/DOCKET NUMBER:  WIS-5-3-84452
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (202)8613000
CC         TELEFAX:  (202)822-9944
CC         TELEX:  6714627 CUSH
CC     INFORMATION FOR SEQ ID NO:  5:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  795 amino acids
CC         TYPE:  AMINO ACID
CC         TOPOLOGY:  linear
SQ   SEQUENCE   795 AA; 87498 MW; 3326890 CN;

  Query Match             16.3%;  Score 577;  DB 3;  Length 795;
  Best Local Similarity  32.9%;  Pred. No. 4.15e-41;
  Matches    92;  Conservative   83;  Mismatches  89;  Indels  16;  Gaps  15;

Db     424 SPPSSSSAATGPPPKICLVCSDEASGCHYGVLTCGSCKVFFKRAVEGQHNYLCAGRNDCI 483
            ::| :::::       . .| |||| ||| ||||| .| .|| ||||  .:.:.| |    |.:
Qy      81 ASPVIGPGSK-RDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCT 139

Db     484 IDKIRRKNCPACRYRKCLQAGMNL-EARKTK-K-KIKGIQQATAG-V-SQDTSENPN-KT 537
            ||| |||:| ||: ||| : ||       .|    .      .     :
Qy     140 IDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHA 199

Db     538 -IVPAALPQ-LTPT-IV-SLLEVIEPEVLYAGYDSSVPDSAWRIMTTLNMLGGRQVIAAV 593
              |    | : |:|  || .|||    | ||  .    ..    ..      | .|  |     ..
Qy     200 PRVRELLLDALSPEQLVLTLLEAEPPHVLISRPSAPFIEAS--MMMSLIKLADKELVHMI 257

Db     594 KWAKAILGLPNLHLDDQMTLLQYSWMFLMAFALGWRSYRQSSGNLLCFAPDLIIN-EQRM 652
             ||| | |: :| | ||: ||:  || ::  :| |||     .  .  .    .....
Qy     258 SWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHP-GKLI-FAPDLVLDRDEGK 315

Db     653 SIPCMYDQCKHMLFVSSELQRLQVSYEEYLCMKTLLLLSS 692
             :  :  :      :|  :| :: |  :      ||||:|  . .: |
Qy     316 CVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNS 355

RESULT     5
ID   US-07-952-800-2      STANDARD;      PRT;    533 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 2, Application US/07952800
XX
CC    Sequence 2, Application US/07952800
CC    Patent No. 5403925
CC     GENERAL INFORMATION:
CC       APPLICANT:  OZATO, KEIKO
CC       TITLE OF INVENTION:  A NEW MEMBER OF THE NUCLEAR HORMONE
CC       TITLE OF INVENTION:  RECEPTOR SUPERFAMILY AND A CDNA CLONE THEREOF
CC       NUMBER OF SEQUENCES:  4
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  TOWNSEND AND TOWNSEND
CC         STREET:  1 MARKET PLAZA  STEUART TOWER  SUITE 2000
CC         CITY:  SAN FRANCISCO
CC         STATE:  CA
CC         COUNTRY:  USA
CC         ZIP:  94105
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/952,800
CC         FILING DATE:  19920928
CC         CLASSIFICATION:  435
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Weber, Kenneth A
CC         REGISTRATION NUMBER:  31,677
```

```
CC         REFERENCE/DOCKET NUMBER:  15280-21-1
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  415-543-9600
CC         TELEFAX:  415-543-5043
CC     INFORMATION FOR SEQ ID NO:  2:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  533 amino acids
CC         TYPE:  AMINO ACID
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   533 AA; 56921 MW; 1406341 CN;

  Query Match             15.1%;  Score 533;  DB 5;  Length 533;
  Best Local Similarity  34.1%;  Pred. No. 4.13e-37;
  Matches   128;  Conservative   85;  Mismatches 132;  Indels  30;  Gaps  21;

Db     186 VRGLHCPPPPGGPGAGK--RICAICGDRSSGKHYGVYSCEGCKGFFKRTIRKDLIYSCRD 243
            | | .|..|  |||. .   ::|||:|:| :|| |||| ||||||:||||:|:      | |
Qy      75 VSGNPCASPVIGPGSKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPA 134

Db     244 NKDCIVDKRQRNRCQYCRYQKCLATGM-K----RE--AVQ-EERQRGKDKDGDGEG-AGG 294
             ..||.|| :|. || || .||   || |    ||   . .   |||: | :     | |
Qy     135 NQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKR 194

Db     295 APEEMPVDR-ILEAELAVEQKS-DQ-GVEGPGGIGGSGSSP-NDP--VINICQAADKQLF 348
            :    |   | .|    |: ||           | |    :  |:|  ::  : ::   |||:|
Qy     195 SGGHAPPVPELLLDALSPEQLVLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELV 254

Db     349 TLVEWAKRIPHFSSLPLDDQVILLRAGWNELLIASFSHRSIDVRDGILLAIGLHVHRNSA 408
             .: |||.|| |  |.| ||| || . | |.|: ..  ||||    ..:|  | : |: :
Qy     255 HMIFWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEG 314

Db     409 HSA-GVGAIFDRVLTELVSKMRDMRMDKTELGCLRAIILFNPDAKGLSNPS-EVEV---L 463
               |.  |||  .:.    |. |.....  |  |..|.||:|:     |  :: : :
Qy     315 KCVEGILEIFDMLLAT-TSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKL 373

Db     464 -R--EKVYASLE-TYCKQKYP-EQQG-RFAKLLLRLPALRSIGLKCLEHLFFFKLIGDTP 517
             :   . |   :|       |     . :||. |.|.||. |. :|   : | :|||:   |     |
Qy     374 AHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVP 433

Db     518 IDTFLMEMLEAPHQL 532
            :   :|:|||:| | |
Qy     434 VYDLLLEMLNA-HVL 447

RESULT     6
ID   US-07-952-800-4      STANDARD;      PRT;    446 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 4, Application US/07952800.
XX
CC    Sequence 4, Application US/07952800
CC    Patent No. 5403925
CC     GENERAL INFORMATION:
CC       APPLICANT:  OZATO, KEIKO
CC       TITLE OF INVENTION:  A NEW MEMBER OF THE NUCLEAR HORMONE
CC       TITLE OF INVENTION:  RECEPTOR SUPERFAMILY AND A CDNA CLONE THEREOF
CC       NUMBER OF SEQUENCES:  4
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  TOWNSEND AND TOWNSEND
CC         STREET:  1 MARKET PLAZA, STEUART TOWER, SUITE 2000
CC         CITY:  SAN FRANCISCO
CC         STATE:  CA
CC         COUNTRY:  USA
CC         ZIP:  94105
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC        CURRENT APPLICATION DATA:
CC          APPLICATION NUMBER   [illegible]
CC          FILING DATE:  19920928
CC          CLASSIFICATION:  435
CC        ATTORNEY/AGENT INFORMATION:
CC          NAME:  Weber, Kenneth A
CC          REGISTRATION NUMBER:  31,677
CC          REFERENCE/DOCKET NUMBER:  15280-21-1
CC        TELECOMMUNICATION INFORMATION:
CC          TELEPHONE:  415-543-9600
CC          TELEFAX:  415-543-5043
CC      INFORMATION FOR SEQ ID NO.  4.
CC        SEQUENCE CHARACTERISTICS:
CC          LENGTH:  446 amino acids
CC          TYPE:  AMINO ACID
CC          TOPOLOGY:  linear
CC        MOLECULE TYPE:  protein
SQ   SEQUENCE   446 AA; 48023 MW; 964846 CN;

  Query Match              15.0%;  Score 529;  DB 5;  Length 446;
  est Local Similarity  34.1%;  Pred. No. 9.51e-37;
  atches   128;  Conservative   84;  Mismatches 133;  Indels  30;  Gaps  21;

Db      99 VRGLHCPPPPGGPGAGK--RLCAICGDRSSGKHYGVYSCEGCKGFFKRTIRKDLIYSCRD 156
            | |  :|::|  |||: :   ::||:|:| :|| |||| |||||||:||||:|:      | |
Qy      75 VSGNRCASPVGGPGSKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPA 134

Db     157 NKDCTVDKRQRNRCQYCRYQKCLATGM-K----PE--AVQ-EERQRGKDKDGDGDG-AGG 207
             ::||:|| :|: || || :||     || |      ||  : :    |||: | :      | |
Qy     135 TNQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRREPCGYRLVRRQRSADEQLHCAGKAKP 194

Db     208 APEEMPVDR-ILEAELAVEQKS-DQ-GVEGPGATGGGGSSP-NDP--VTNICQAADKQLF 261
            :     |   | :|    |: ||                 | |     :  |:|  ::   : ::    |||:|
Qy     195 SGGHAPRVRELLLDALSPEQLVLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELV 254

Db     262 TLVEWAKRIPHFSSLPLDDQVILLRAGWNELLIASFSHRSIDVRDGILLATGLHVHRNSA 321
             :: |||:||  |   |:| ||| || : | |:|: ::  ||||     :::|  | : |: :
Qy     255 HMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEG 314

Db     322 HSA-GVGAIFDRVLTELVSKMRDMRMDKTELGCLRAIILFNPDAKGL-SNPGEVEI---L 376
                |:   ||| :|:     |:  |:::::   |    |::|:||:|:       |  :    :  :     |
Qy     315 KCVEGILEIFDMLLAT-TSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKL 373

Db     377 -R--EKVYASLE-TYCKQKYP-EQQG-RFAKLLLRLPALRSIGLKCLEHLFFFKLIGDTP 430
             :  : |   :|       |     :  :|||: |:|:||: |:  :|   : | :|||:   |      |
Qy     374 AHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVP 433

Db     431 IDTFLMEMLEAPHQL 445
            :   :|:|||:|  | |
    434 VYDLLLEMLNA-HVL 447

RESULT      7
ID    PCT-US91-00399-4      STANDARD;      PRT;    467 AA.
XX
AC    xxxxxx
XX
DT    01-JAN-1900
XX
DE    Sequence 4, Application PC/TUS9100399.
XX
CC     Sequence 4, Application PC/TUS9100399
CC      GENERAL INFORMATION:
CC        APPLICANT:  Mangelsdorf, Dr., David J.
CC        APPLICANT:  Evans Dr., Ronald M.
CC        TITLE OF INVENTION:  RETINOID RECEPTOR COMPOSITIONS AND
CC        TITLE OF INVENTION:  METHODS
CC        NUMBER OF SEQUENCES:  7
CC        CORRESPONDENCE ADDRESS:
CC          ADDRESSEE:  FITCH, EVEN, TABIN & FLANNERY
CC          STREET.  135 South LaSalle Street, Suite 900
CC          CITY:  Chicago
CC          STATE:  Illinois
```

```
CC          COUNTRY:  USA
            ZIP   [illegible]
CC        COMPUTER READABLE FORM:
CC          MEDIUM TYPE:  Floppy disk
CC          COMPUTER:  IBM PC compatible
CC          OPERATING SYSTEM:  PC-DOS/MS-DOS
CC          SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC        CURRENT APPLICATION DATA:
CC          APPLICATION NUMBER:  PCT/US91/00399
CC          FILING DATE:  19910122
CC          CLASSIFICATION:  435
CC        PRIOR APPLICATION DATA:
CC          APPLICATION NUMBER:  07/478,071
CC          FILING DATE:  09-FEB-1990
CC        ATTORNEY/AGENT INFORMATION:
CC          NAME:  Reiter, Mr., Stephen E.
CC          REGISTRATION NUMBER:  31192
CC          REFERENCE/DOCKET NUMBER:  50852
CC        TELECOMMUNICATION INFORMATION:
CC          TELEPHONE:  (619) 552-1311
CC          TELEFAX:  (619) 552-0095
CC          TELEX.  20 6566 PAILAW CGO
CC      INFORMATION FOR SEQ ID NO:  4:
CC        SEQUENCE CHARACTERISTICS:
CC          LENGTH:  467 amino acids
CC          TYPE:  AMINO ACID
CC          TOPOLOGY:  linear
CC        MOLECULE TYPE:  protein
SQ   SEQUENCE   467 AA; 51216 MW; 1083042 CN;

  Query Match              14.6%;  Score 514;  DB 9;  Length 467;
  Best Local Similarity  33.0%;  Pred. No. 2.17e-35;
  Matches   116;  Conservative   89;  Mismatches 119;  Indels  28;  Gaps  21;

Db     138 HICAICGDRSSGKHYGVYSCEGCKGFFKRTVRKDLIYPCRDNKDCLIDKRQRNRCQYCRY 197
            |.||.|.| .|| |||| ||||||.||||.::       | |     ::| ||| :|: || ||
Qy      94 HFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRL 153

Db     198 QKCLAMGMKREAVQEER-Q-R-GK-DRNENE-VESTSSAN---EDMP-VEKILFAELAVE 248
            .||   .||  .  .  .  ||     |   .  .|.  .|  .  .: |:          | |  :|   |: |
Qy     154 RKCYEVGMVKCGSRREPCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPE 213

Db     249 PKTETYVEANMG--L-N-PSSP-NDP--VTNICQAADKQLFTLVEWAKRIPHFSELPLDD 301
                 | :||:      | : ||:|  ::   : ::    |||:|   :: |||:|| | ||:| |
Qy     214 QLVITLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 273

Db     302 QVILLRAGWNFLLIASFSHRSIAVKDGILLATGLHVHRNSAHSA-GVGAIFDRVLTELVS 360
            || || : | |:|: ::  |||       :::|  | : |: :       |:  ||| :|:      |
Qy     274 QVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVIDRDEGKCVEGILEIFDMLLAT-TS 332

Db     361 KMRDMQMDKTELGCLRAIVLFNPDSKGLSNPA-EVEA---L-R--EKVYASLE-AYCKHK 412
            : |:: ::   |   |::|::|:|:       |  :: : ::      |  :   : |    :|       |
Qy     333 RFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKSG 392

Db     413 YP-EQPG-RFAKLLLRLPALRSIGLKCLEHLFFFKLIGDTPIDTFLMEMLEA 462
             : :| : |:|:||: |: :|   : | :|||:  |       |:  :|:|||:|
Qy     393 ISSQQQSMRLANLLMLLSHVPHASNKGMEHLLNMKCKNVVPVYDLLLEMLNA 444

RESULT      8
ID    PCT-US91-00399-2      STANDARD;      PRT;    462 AA.
XX
AC    xxxxxx
XX
DT    01-JAN-1900
XX
DE    Sequence 2, Application PC/TUS9100399.
XX
CC     Sequence 2, Application PC/TUS9100399
CC      GENERAL INFORMATION:
CC        APPLICANT:  Mangelsdorf, Dr., David J.
CC        APPLICANT:  Evans Dr., Ronald M.
CC        TITLE OF INVENTION:  RETINOID RECEPTOR COMPOSITIONS AND
CC        TITLE OF INVENTION:  METHODS
CC        NUMBER OF SEQUENCES:  7
CC        CORRESPONDENCE ADDRESS:
CC          ADDRESSEE:  FITCH, EVEN, TABIN & FLANNERY
CC          STREET:  135 South LaSalle Street, Suite 900
CC          CITY:  Chicago
CC          STATE:  Illinois
CC          COUNTRY:  USA
CC          ZIP:  60603
CC        COMPUTER READABLE FORM:
CC          MEDIUM TYPE:  Floppy disk
CC          COMPUTER:  IBM PC compatible
CC          OPERATING SYSTEM:  PC-DOS/MS-DOS
CC          SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC        CURRENT APPLICATION DATA:
CC          APPLICATION NUMBER:  PCT/US91/00399
CC          FILING DATE:  19910122
CC          CLASSIFICATION:  435
CC        PRIOR APPLICATION DATA:
            APPLICATION NUMBER:  07/478,071
            FILING DATE:  09-FEB-1990
          ATTORNEY/AGENT INFORMATION:
CC          NAME:  Reiter, Mr., Stephen E
CC          REGISTRATION NUMBER:  31192
CC          REFERENCE/DOCKET NUMBER:  50852
CC        TELECOMMUNICATION INFORMATION:
CC          TELEPHONE:  (619) 552-1311
CC          TELEFAX:  (619) 552-0095
CC          TELEX:  20 6566 PATLAW CGO
CC      INFORMATION FOR SEQ ID NO:  2:
CC        SEQUENCE CHARACTERISTICS:
CC          LENGTH:  462 amino acids
CC          TYPE:  AMINO ACID
CC          TOPOLOGY:  linear
CC        MOLECULE TYPE:  protein
SQ   SEQUENCE   462 AA;  50811 MW;  1063517 CN;

  Query Match             14.5%;  Score 513;  DB 9;  Length 462;
  Best Local Similarity   33.0%;  Pred. No. 2.67e-35;
  Matches  116;  Conservative   89;  Mismatches 119;  Indels  28;  Gaps  21;

Db     133 HICAICGDRSSGKHYGVYSCEGCKGFFKRTVRKDLTYTCRDNKDCLIDKRQRNRCQYCRY 192
           |:||:|:| :|| |||| |||||| ||||:       : |     :| ||| ::: || ||
Qy      94 HFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRL 153

Db     193 QKCLAMGMKREAVQEER-Q-R-GK-DRNENE-VESTSSAN---EDMPVER-ILEAELAVE 243
           :||  :|| : : : ||    |  : :|: :| :   :: |:        |  | :|    |: |
       154 RKCYEVGMVKCGSPRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPE 213

       244 PKIETYVEANMG--L-N-PSSP-NDP--VTNICQAADKQLFTLVEWAKRIPHFSELPLDD 296
                | :||:     | : ||:|  ::  :  .   |||.|      |||:|| | ||:| |
Qy     214 QLVLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 273

Db     297 QVILLPAGWNELLIASFSHRSIAVKDGILLATGLHVHRNSAHSA-GVGAIFDRVLTELVS 355
           || ||  :  |  |:|: ::  |||      ...|  | . |: .     |.  ||| .|:   |
Qy     274 QVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLAT-TS 332

Db     356 KMRDMQMDKTELGCLPAIVLFNPDSKGLSNPA-EVEA---L-R--EKVYASLE-AYCKHK 407
           : |:: ::  |  |:::|:: |       :          |      :    :    :
Qy     333 RFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKSG 392

Db     408 YP-EQPG-RFAKLLLRLPALRSIGLKCLEHLFFFKLIGDTPIDTFLMEMLEA 457
            : :| : |.|:|:|. |. .|   . | .|||.  |     |:    : :|| :
Qy     393 ISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNA 444


RESULT    9
ID   PCT-US91-00399-6     STANDARD;     PRT;   463 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
```

```
DE   Sequence 6, Application PC/TUS9100399.
XX
CC    Sequence 6, Application PC/TUS9100399
CC     GENERAL INFORMATION:
CC       APPLICANT:  Mangelsdorf, Dr., David J.
CC       APPLICANT:  Evans Dr., Ronald M.
CC       TITLE OF INVENTION:  RETINOID RECEPTOR COMPOSITIONS AND
CC       TITLE OF INVENTION:  METHODS
CC       NUMBER OF SEQUENCES:  7
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  FITCH, EVEN, TABIN & FLANNERY
CC         STREET:  135 South LaSalle Street, Suite 900
CC         CITY:  Chicago
CC         STATE:  Illinois
CC         COUNTRY:  USA
CC         ZIP:  60603
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US91/00399
CC         FILING DATE:  19910122
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/478,071
CC         FILING DATE:  09-FEB-1990
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Reiter, Mr., Stephen E.
CC         REGISTRATION NUMBER:  31192
CC         REFERENCE/DOCKET NUMBER:  50852
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (619) 552-1311
CC         TELEFAX:  (619) 552-0095
CC         TELEX:  20 6566 PATLAW CGO
CC     INFORMATION FOR SEQ ID NO:  6:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  463 amino acids
CC         TYPE:  AMINO ACID
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   463 AA;  50978 MW;  1097932 CN;

  Query Match             13.8%;  Score 486;  DB 9;  Length 463;
  Best Local Similarity   32.2%;  Pred. No. 7.35e-33;
  Matches  117;  Conservative   94;  Mismatches 120;  Indels  32;  Gaps  24;

Db     127 PSTSPGSLVK-HICAICGDRSSGKHYGVYSCEGCKGFFKRTIRKDLIYTCRDNKDCLIDK 185
           | |.|||    |:||:|:| :|| |||| ||||||:||||:|:      | |    ::| |||
Qy      83 PVTGPGSKPDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDK 142

Db     186 RQRNRCQYCRYQKCLVMGMKREAVQEER-Q-R--SRER-AESEAECASSS-HED--MP-V 236
            :|: || || :||  :|| : : : ||    |    |:| |: :  ||: : :      | |
Qy     143 NRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRV 202

Db     237 -ERIIEA---E-LAVEP-KTES-YCDMNVENS--INDPVI-NICHAADKQLFILVEWAKR 286
             | :|:|   | | :      :|:      ::   ::  |: ::   ::   |||:|  :: |||:
Qy     203 RELLLDALSPEQLVLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKK 262

Db     287 IPHFSDLTLEDQVILLRAGWNELLIASFSHRSVSVQDGILLATGLHVHRSSAHS-RGVGS 345
           || | :|.| ||| || . | |.|. ::  ||:       :::|  | : |  :       |:
Qy     263 IPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILE 322

Db     346 IFDRVLTELVSKMKDMQMDKSELGCLRAIVLFNPDAKGLSNPS-EVET---L-R--EKVY 398
           ||| :|:    |: ::: ::   |   |::|::|:|:     |  :: : ::    | :  : |
Qy     323 IFDMLLAT-TSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVT 381

Db     399 ATLF-AYTKQKYP-EQPG-RFAKLLLRLPALRSIGLKCLEHLFFFKLIGDTPIDSFLMEM 455
            :|     :|    : :| : |:|:||: |: :|   : | :|||:   |       |:  :|:||
Qy     382 DALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEM 441

Db     456 LET 458
```

```
           |::
Qy     442 LNA 444


RESULT   10
ID   5260432-2     STANDARD;     PRT;   492 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Patent No. 5260432.
XX
CC   Patent No. 5260432
CC       APPLICANT: TAKAKU, FUMIMARO;ISHIKAWA, TAKASHI;IMAWARI, MICHIO;
CC   EVANS, RONALD M.;UMESONO, KAZUHIKO
CC       TITLE OF INVENTION: HUMAN GAMMA RETINOIC ACID RECEPTOR DNA
CC       NUMBER OF SEQUENCES: 4
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER: US/07/370,407
CC         FILING DATE: 22 JUN 1989
CC   SEQ ID NO:2:
CC         LENGTH: 454
SQ   SEQUENCE   492 AA; 54527 MW; 1333563 CN;

  Query Match              12.5%;  Score 441;  DB 1;  Length 454;
  Best Local Similarity   28.4%;  Pred. No. 8.17e-29;
  Matches   99;  Conservative   88;  Mismatches 140;  Indels   22;  Gaps   18;

Db      90 CFVCNDKSSGYHYGVSSCEGCKGFFRRSIQKNMVYTCHRDKNCIINKVTRNRCQYCRLQK 149
           | ||:| :||||||| ||||||:||:|||| :   | |    : | |:|  |: || |||:|
Qy      96 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRLRK 155

Db     150 CFEVGMSKEAVRNDR-N-K-KKEVKEEGSPDSYELSPQLEELITKVSKAHQETF-PS-L 204
           |:|||| | : | :|   :  :::   :               : :        :|      ::: |  |
Qy     156 CYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 215

Db     205 C-QL-GKYTTNSSADHR-VQL-DLGLWDKFSELATKCIIKIVEFAKRLPGFTGLSIADQI 260
              |     :     :     : : ::    :: || | :: ::  ||::|||  ||: ||:
Qy     216 VLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFDQV 275

Db     261 TLLKAAC-LDILMLRICTRYTPEQDTMTFSDGLTLNRTQ-MHNAGFGPLTDLVFAFAGQL 318
            ||   :| :::||: :  |            : |:  | |:| : :    |:  : |:::| ::::
Qy     276 RLL-ESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSRF 334

Db     319 LPLEMDDTE----TG--LL-SAICLICGDRMDLEEPEKVDKLQEPLLEAL-RLYARRRRP 370
            | ::  |       :  || |::    :       | : : |:  | ::: :|| :: |:    :
Qy     335 RELKLQHKEYLCVKAMILLNSSMYPLVIATQDADSSRKLAHLLNAVTDALVWVIAKSGIS 394

Db     371 SQP-YM-FPRMLMKITDLRGISTKGAERAITLKMEIPGPMPPLIREMLE 417
           ||  | :: :|| :: :|  | || |: : :|       |:  |: |||
Qy     395 SQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLN 443


RESULT   11
ID   5223606-3     STANDARD;     PRT;   398 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Patent No. 5223606.
XX
CC   Patent No. 5223606
CC       APPLICANT:  BLAUDIN DE THE, HUGHES;MARCHIO, AGNES;TIOLLAIS,
CC   PIERRE;DEJEAN, ANNE
CC       TITLE OF INVENTION: STEROID/THYROID HORMONE RECEPTOR-RELATED
CC   PROTEIN INAPPROPRIATELY EXPRESSED IN HUMAN HEPATOCELLULAR CARCINOMA
CC       NUMBER OF SEQUENCES: 11
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER: US/07/134,130
CC         FILING DATE: 17-DEC-1987
CC       PRIOR APPLICATION DATA:
CC   SEQ ID NO:3:
CC         LENGTH: 368
SQ   SEQUENCE   398 AA; 44842 MW; 907111 CN;

  Query Match              12.4%;  Score 439;  DB 1;  Length 368;
  Best Local Similarity   26.9%;  Pred. No. 1.23e-28;
  Matches   94;  Conservative   97;  Mismatches 136;  Indels   22;  Gaps   18;

Db       1 CFVCQDKSSGYHYGVSACEGCKGFFRRSIQKNMIYTCHRDKNCVINKVTRNRCQYCRLQK 60
           | || | :||||||| :|||||:||:|||| :   | |    : | |:|  |: || |||:|
Qy      96 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRLRK 155

Db      61 CFEVGMSKESVRNDR-N-K-KKETSKQECTESYEMTAELDDLTEKIRKAHQETF-PS-L 115
           |:|||| | : | :|   :  ::: | :|         ::          : ::|      ::: |  |
Qy     156 CYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 215

Db     116 C-QL-GKYTINSSADHR-VRL-DLGLWDKFSELATKCIIKIVEFAKRLPGFTGLTIADQI 171
              |     :     :     : : ::    :: || | :: ::  ||::|||  |:: ||:
Qy     216 VLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFDQV 275

Db     172 TLLKAAC-LDILILRICTRYTPEQDTMTFSDGLTLNRTQ-MHNAGFGPLTDLVFTFANQL 229
            ||   :| :::|::: :  |            : |:  | |:| : :    |:  : |:::: ::::
Qy     276 RLL-ESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSRF 334

Db     230 LPLEMDDTE----TG--LL-SAICLICGDRQDLEEPTKVDKLQEPLLEALK-IYIRKRRP 281
            | ::  |       :  || |::    :       || : : |:  | ::: :||  :  :    :
Qy     335 RELKLQHKEYLCVKAMILLNSSMYPLVIATQDADSSRKLAHLLNAVTDALVWVIAKSGIS 394

Db     282 SKPH-M-FPKILMKITDLRSISAKGAERVITLKMEIPGSMPPLIQEMME 328
           |  : | ::::|| :: :|  | || |::: :|       ::  |: ||::
Qy     395 SQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLN 443


RESULT   12
ID   5223606-2     STANDARD;     PRT;   485 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Patent No. 5223606.
XX
CC   Patent No. 5223606
CC       APPLICANT:  BLAUDIN DE THE, HUGHES;MARCHIO, AGNES;TIOLLAIS,
CC   PIERRE;DEJEAN, ANNE
CC       TITLE OF INVENTION: STEROID/THYROID HORMONE RECEPTOR-RELATED
CC   PROTEIN INAPPROPRIATELY EXPRESSED IN HUMAN HEPATOCELLULAR CARCINOMA
CC       NUMBER OF SEQUENCES: 11
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER: US/07/134,130
CC         FILING DATE: 17-DEC-1987
CC       PRIOR APPLICATION DATA:
CC   SEQ ID NO:2:
CC         LENGTH: 448
SQ   SEQUENCE   485 AA; 54437 MW; 1345912 CN;

  Query Match              12.4%;  Score 439;  DB 1;  Length 448;
  Best Local Similarity   26.9%;  Pred. No. 1.23e-28;
  Matches   94;  Conservative   97;  Mismatches 136;  Indels   22;  Gaps   18;

Db      81 CFVCQDKSSGYHYGVSACEGCKGFFRRSIQKNMIYTCHRDKNCVINKVTRNRCQYCRLQK 140
           | || | :||||||| :|||||:||:|||| :   | |    : | |:|  |: || |||:|
Qy      96 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRLRK 155

Db     141 CFEVGMSKESVRNDR-N-K-KKETSKQECTESYEMTAELDDLTEKIRKAHQETF-PS-L 195
           |:|||| | : | :|   :  ::: | :|         ::          : ::|      ::: |  |
Qy     156 CYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 215

Db     196 C-QL-GKYTINSSADHR-VRL-DLGLWDKFSELATKCIIKIVEFAKRLPGFTGLTIADQI 251
              |     :     :     : : ::    :: || | :: ::  ||::|||  |:: ||:
Qy     216 VLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFDQV 275

Db     252 TLLKAAC-LDILILRICTRYTPEQDTMTFSDGLTLNRTQ-MHNAGFGPLTDLVFTFANQL 309
```

```
              ||  :| :::|:: :  |          : |:  | |.|  . :   |.  . |.... ....
Qy     276 PLL-ESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLIATTSRF 334

Db     310 LPLEMDDTE---TG--LL-SAICLICGDRQDLEEPTKVDKLQEPLLEALK-IYIRKRRP 361
               | ::  |       :  || |::   :      || : : |:  | ::: :||   :   :    :
Qy     335 RELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKSGIS 394

Db     362 SRPH-M-FPKIIMKITDLRSISAKGAERVITLKMEIPGSMPPLIQEMME 408
             :  : | ::::|| :: :|  | || |::: :|       ::  |: ||::
Qy     395 SQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLN 443


RESULT   13
ID   PCT-US92-02320A-4     STANDARD      PRT    462 AA
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
     Sequence 4, Application PC/TUS9202320A
XX
     Sequence 4, Application PC/TUS9202320A
CC    GENERAL INFORMATION:
CC       APPLICANT:  Sloan-Kettering Institute, For Cancer Research
CC       TITLE OF INVENTION:  METHODS FOR DETECTION AND TREATMENT OF CANCER
CC       NUMBER OF SEQUENCES:  6
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  COOPER & DUNHAM
CC         STREET:  30 ROCKEFELLER PLAZA
CC         CITY:  NEW YORK
CC         STATE:  NEW YORK
CC         COUNTRY:  U.S.A.
CC         ZIP:  10112
CC       COMPUTER READABLE FORM
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.24
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US92/02320A
CC         FILING DATE:
CC         CLASSIFICATION:
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 675,084
CC         FILING DATE:  22-MAR-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 673,838
           FILING DATE:  22-MAR-1991
         ATTORNEY/AGENT INFORMATION:
CC         NAME:  WHITE, JOHN P
CC         REGISTRATION NUMBER:  28,678
CC         REFERENCE/DOCKET NUMBER:  38694-PCT
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 977-9550
CC         TELEFAX:  (212) 644-0525
CC         TELEX:  (212) 422523 COOP UI
CC     INFORMATION FOR SEQ ID NO:  4:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  462 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   462 AA  50771 MW  1101259 CN;

  Query Match            12.1%,  Score 428,  DB 10,  Length 462,
  Best Local Similarity  27.9%   Pred. No. 1.19e-27,
  Matches    97,  Conservative   92,  Mismatches 138,  Indels  22,  Gaps  20,

Db      88 CFVCQDKSSGYHYGVSACEGCKGFFRRSIQKNMVYTCHRDKNCIINKVTRNRCQYCRLQK 147
             |  || |  :||.|||| :||||||:||:||||  :    | |     . |  |:|   |. || |||.|
Qy      96 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRK 155

Db     148 CFEVGMSKESVRNDR-N-K-KKKEVPKPECSESYTLTPEVGELIEKVRKAHQETF-PA-L 202
```

```
             |:||||  |  :  |  :|    :   ::: :  |          :    |     :||     ::: |  |
Qy     156 IYEVDMVKIDSHHHSIDYRLVRRLPSALEQLHQAGKAKRSGGHAPRVRLLLLLALSRLQL 215

Db     203 C-QL-GFYTTNN-SSEQRVSL-DLLLWLKFSELSIKGIIKIVEFAKQLPGFIILIIADQI 258
             |      :   |      :: :  :     :: |: | ::   :  || :|||  |:: ||:
Qy     216 VLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFDQV 275

Db     259 TLLKAAC-IDILILRICTRYTPEQDTMIFSDGLTLNRIQ-MHNAGFGPLTDLVFAFANQL 316
             ||  :| :::|:: :  |          : |:  | |:| : :    |:  : |:::| ::::
Qy     276 PLL-ESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLIATTSRF 334

Db     317 --LPLE-MDD--AETG-LL-SAICLICGDRQDLEQPDRVDMLQEPLLEALK-VYVRKRRP 368
               | |. ..      .  || ||.   :      ||  .  .  :.  |  . . .||  |   .   .
Qy     335 RELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKSGIS 394

Db     369 SRPH-M-FPKMLMKITDLRSISAKGAERVITLKMEIPGSMPPLIQEMLE 415
             |: : | ::::|| :: :|  | || |::: :|      ::  |: |||:
Qy     395 SQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLN 443


RESULT   14
ID   PCT-US92-02320A-2     STANDARD;     PRT;    797 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 2, Application PC/TUS9202320A.
XX
CC    Sequence 2, Application PC/TUS9202320A
CC     GENERAL INFORMATION:
CC       APPLICANT:  Sloan-Kettering Institute, For Cancer Research
CC       TITLE OF INVENTION:  METHODS FOR DETECTION AND TREATMENT OF CANCER
CC       NUMBER OF SEQUENCES:  6
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  COOPER & DUNHAM
CC         STREET:  30 ROCKEFELLER PLAZA
CC         CITY:  NEW YORK
CC         STATE:  NEW YORK
CC         COUNTRY:  U.S.A.
CC         ZIP:  10112
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.24
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US92/02320A
CC         FILING DATE:
CC         CLASSIFICATION:
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 675,084
CC         FILING DATE:  22-MAR-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 673,838
CC         FILING DATE:  22-MAR-1991
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  WHITE, JOHN P
CC         REGISTRATION NUMBER:  28,678
CC         REFERENCE/DOCKET NUMBER:  38694-PCT
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 977-9550
CC         TELEFAX:  (212) 644-0525
CC         TELEX:  (212) 422523 COOP UI
CC     INFORMATION FOR SEQ ID NO:  2:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  797 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   797 AA;  89281 MW;  3152972 CN;

  Query Match            12.1%;  Score 428;  DB 10;  Length 797;
  Best Local Similarity  27.8%;  Pred. No. 1.19e-27;
  Matches   97;  Conservative   92;  Mismatches [illegible];  Indels  22;  Gaps  [illegible];

Db      423 CFVCQDKSSGYHYGVSACEGCKGFFRRSIQKNMVYTCHRDKNCIINKVTRNRCQYCRLQK 482
             | ||  |  :|||||||  :|||||:||:||||  :    | |     :  | |:|   |: || |||:|
Qy       96 CAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRLRK 155

Db      483 CFEVGMSKESVRNDR-N-K-KKKEVPKPECSESYTLTPEVGELIEKVRKAHQETF-PA-L 537
             |:||||  |  :  |  :|      :    :::  :   |           :    |      :||      ::: |   |
Qy      156 CYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPEQL 215

Db      538 C-QL-GKYTTNN-SSEQRVSL-DIDLWDKFSELSTKCIIKTVEFAKQLPGFTTLTIADQI 593
                |           :     |        ::  :    :       :: |: |  ::    :    || ·|||  |·· ||·
Qy      216 VLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFDQV 275

Db      594 TLLKAAC-LDILILRICTRYTPEQDTMTFSDGLTLNRTQ-MHNAGFGPLTDLVFAFANQL 651
             ||   :|  :::|::  :    |                :  |:    | |:|  :  :     |:    :  |:::|  ::::
Qy      276 RLL-ESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSRF 334

Db      652 --LPLE-MDD--AETG-LL-SAICLICGDRQDLEQPDRVDMLQEPLLEALK-VYVRKRRP 703
               | |: ::         :   || |::    :        ||  :   :  ::    |  :::  :||     |     :      :
Qy      335 RELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKSGIS 394

Db      704 SRPH-M-FPKMLMKITDLRSISAKGAERVITLKMEIPGSMPPLIQEMLE 750
             |:  :  |  ::::||  ::  :|     |  ||  |:::  :|           ::    |:  |||:
Qy      395 SQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLN 443

RESULT    15
ID   US-08-333-358-8      STANDARD;      PRT;   440 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 8, Application US/08333358.
XX
CC    Sequence 8, Application US/08333358
CC    Patent No. 5571696
CC     GENERAL INFORMATION:
CC        APPLICANT:  EVANS Ph.D., RONALD M.
CC        APPLICANT:  MANGELSDORF Ph.D., DAVID J.
CC        APPLICANT:  ONG Ms., ESTELITA S.
CC        APPLICANT:  ORO Ph.D., ANTHONY E.
CC        APPLICANT:  BORGMEYER Ph.D., UWE K.
CC        APPLICANT:  GIGUERE Ph.D., VINCENT NMN
CC        APPLICANT:  YAO Mr., TSO-PANG NMN
CC        TITLE OF INVENTION:  NOVEL RECEPTORS
CC        NUMBER OF SEQUENCES:  14
CC        CORRESPONDENCE ADDRESS:
CC          ADDRESSEE:  Pretty, Schroeder, Brueggemann & Clark
CC          STREET:  444 So. Flower St., Suite 2000
CC          CITY:  Los Angeles
CC          STATE:  CA
CC          COUNTRY:  US
CC          ZIP:  90071-2921
CC        COMPUTER READABLE FORM:
CC          MEDIUM TYPE:  Floppy disk
CC          COMPUTER:  IBM PC compatible
CC          OPERATING SYSTEM:  PC-DOS/MS-DOS
CC          SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC        CURRENT APPLICATION DATA:
CC          APPLICATION NUMBER:  US/08/333,358
CC          FILING DATE:
CC          CLASSIFICATION:  435
CC        PRIOR APPLICATION DATA:
CC          APPLICATION NUMBER:  US/07/761,068
CC          FILING DATE:
CC        ATTORNEY/AGENT INFORMATION:
CC          NAME:  Reiter Ph.D., Stephen E.
CC          REGISTRATION NUMBER:  31192
CC          REFERENCE/DOCKET NUMBER:  P31 8936
CC        TELECOMMUNICATION INFORMATION:
CC          TELEPHONE:  (619) 535-9001
CC          TELEFAX:  (619) 535-8949
CC      INFORMATION FOR SEQ ID NO:  8:
CC        SEQUENCE CHARACTERISTICS:
CC          LENGTH:  440 amino acids
CC          TYPE:  amino acid
CC          TOPOLOGY:  linear
CC        MOLECULE TYPE:  protein
SQ   SEQUENCE   440 AA;  49883 MW;  957063 CN;

  Query Match            11.9%;  Score 420;  DB 6;  Length 440;
  Best Local Similarity  30.8%;  Pred. No. 6.18e-27;
  Matches   84;  Conservative   68;  Mismatches 104;  Indels  17;  Gaps  15;

Db       80 GPAPKMLGNELCSVCGDKASGFHYNVLSCEGCKGFFRRSVIKGAHYICHSGGHCPMDTYM 139
             ||::|      :|:||:| |||:|| | ||||||:||:|||:          ||| :   :| :|
Qy       86 GPGSKP-DAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNR 144

Db      140 RRKCQECRLRKCRQAGMREECVLSEEQ--IRLKKLKRQEEEQAHATSLPPRRSSPPQILP 197
             |· || ||||||  · ||     |        |·      || :    |    :|| |  :: :  |  :: :    :
Qy      145 RKSCQACRLRKCYEVGM-VKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVR 203

Db      198 QLSPEQLGMIEKLVAAQQQCNRRSFSDRLRVTPWPMAPDPHSREARQQRFAHFTELAIVS 257
             :|    :  |:    | ||  :   :  :    :  :    |  . .  .|  ::  |  :  ||          :  :|  ||
Qy      204 ELLLDALSP-EQLVLTLLEAE--P-PHVL-IS-RP-SA-PFT-EASM-MMS-LTKLADKE 252

Db      258 VQEIVDFAKQLPGFLQLSREDQIALLKTSAIEVMLLETSRRYNPGSESITFLKDFSYNRE 317
             :    ::     ||  :|||::||    ||: ||  :    :||:::       :|          :     :   |     |:     :|:
Qy      253 LVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRD 312

Db      318 DFAKAGLQV-EFINPIFEFSRAMNELQLNDAEF 349
             :  :|        :  |:::  ::     :           || |        |:
Qy      313 E-GKCVEGILEIFDMLLATTSRFRELKLQHKEY 344

Search completed: Thu Dec 18 08:47:00 1997
Job time : 42 secs.
```

```
  Query Match             18.7%;  Score 662;  DB 7;  Length 930;
  Best Local Similarity   32.5%;  Pred. No. 3.53e-96;
  Matches  106;  Conservative   91;  Mismatches 114;  Indels   15;  Gaps   15;

Db      567 iclicgdeasgchygvltcgsckvffkramegqhnylcagrndcivdkirrkncpacrlr 626
            :| :|:| ||| |||| :| :|| ||||:::|:::|:|:: |:| :|| |||:| |||||
Qy       95 FCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCIIDKNRRKSCQACRLR 154

Db      627 kgqpagrclqgrkfk-kfnkvrvmralldavalpqpvglpnesq-rltfspsqeiq-llp 682
            || : |||   |:|: :  :  ||  |: |    | :  | :: |  :         :: | |
Qy      155 KCYEVGMVKCGSRRERCGYRIVRRQRSADE-QL-HCAGKAKRSGGHAPRVRELLLDALSP 212

Db      683 -plinllmsiepdviyaqhdntkpdtsssllltslnqlgerqllsvvkwskslpgfrnlhi 741
              |:  |:  ||  :  .. .. | | .|.. ||  |....|. .. |.| :||| :| :
Qy      213 FQLVITLIEAFPPHVLISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSL 271

Db      742 ddqitliqyswmslmvfqlgwrsykhvsgqmlyfapdlilneqrmk-essfyslcltmwq 800
             ||: |::  || :.. || |||  | .| .. |||||.|. .  |    ::  :   :
Qy      272 FDQVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDMLLA 329

        801 ipqetvklqvsqeeflcmkvllllnt i-pleglrsqsqfeemrssyirelikaig-lrq 857
               |  |  :  :  |:||:|  ::|||:  :  ||      :::    :  :  :  :  |:  :
        330 TTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIA 389

Db      858 kgvvsssqrfyqltklldnlhdlvkq 883
            |: :|| |:  :|::||  |   |::
Qy      390 KSGISSQQQSMRLANLL-MLLSHVRH 414

Search completed: Thu Dec 18 08:44:42 1997
Job time : 107 secs.
```

```
Db    616 kcrqagmvl-ggrkfk-kfnkvrvmraldgvalpqsvafpnesqtlsqritfspnqeiql 673
          |  :  |||  |:|: :  :  ||  |: |   | :  : :: |   : |:      :  |
Qy    155 KCYEVGMVKGGSRKERCGYRLVRRQRSADE-QL-HCAGKAKRSGGHAPRVR-ELLLDA-L 210

Db    674 vp-plinllmsiepdvvyaghdntkpdtssslltslnqlgerqllsvvkwskslpgfrnl 732
           |  |:  |:  ||  |          | |  |  ||  |    |:     : :  .||: .|
Qy    211 SPEQLVLTLLEAEPPHVLISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVEL 269

Db    733 hiddqitliqyswmslmvfglgwrsykhvsgqmlyfapdliln-eq-r-mke-lsfyslc 788
           :  |: |::  || :::  || |||   | :| :: ||||:|: :: :  :   | :: :
Qy    270 SLFDQVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDM- 326

Db    789 ltmwqipqefvklqvtheeflcmkvllllnt-i-pleglrsqsqfeemrssyirelikai 846
           . :        |  | : | |:||:| ::|||: : ||     :::   : :  :  :  |:
Qy    327 L-L-ATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDAL 384

Db    847 g-lrqkgvvpssqrfyqltklldslhdlvkq 876
           :  |: ::| |:  :|::||  |   |::
Qy    385 VWVIAKSGISSQQQSMRLANLL-MLLSHVRH 414


RESULT   14
ENTRY              QRHUP      #type complete
TITLE              progesterone receptor form B - human
ALTERNATE_NAMES    hPP
CONTAINS           progesterone receptor form A
ORGANISM           #formal_name Homo sapiens #common_name man
DATE               30-Jun-1987 #sequence_revision 18-Nov-1994 #text_change
                     15-Nov-1996
ACCESSIONS         S09971; S12464; A03245
REFERENCE          S09971
   #authors        Kastner, P.; Krust, A.; Turcotte, B.; Stropp, U.; Tora, L.;
                     Gronemeyer, H.; Chambon, P.
   #journal        EMBO J. (1990) 9:1603-1614
   #title          Two distinct estrogen-regulated promoters generate
                     transcripts encoding the two functionally different human
                     progesterone receptor forms A and B.
   #cross-references MUID:90228361
   #accession      S09971
      ##molecule_type mRNA
      ##residues      1-933 ##label KAS
      ##cross-references EMBL:X51730
REFERENCE          S12464
   #authors        Kastner, P.
   #submission     submitted to the EMBL Data Library, February 1990
   #accession      S12464
      ##molecule_type mRNA
      ##residues      1-343,'T',345-933 ##label KA2
      ##cross-references EMBL:X51730
REFERENCE          A03245
   #authors        Misrahi, M.; Atger, M.; D'Auriol, L.; Loosfelt, H.; Meriel,
                     C.; Fridlansky, F.; Guiochon-Mantel, A.; Galibert, F.;
                     Milgrom, E.
   #journal        Biochem. Biophys. Res. Commun. (1987) 143:740-748
   #title          Complete amino acid sequence of the human progesterone
                     receptor deduced from cloned cDNA.
   #cross-references MUID:87184565
   #accession      A03245
      ##molecule_type mRNA
      ##residues      1-225,'G',227-255,'V',257-659,'V',661-933 ##label MIS
      ##cross-references GB:M15716
GENETICS
   #gene           GDB:PGR
      ##cross-references GDB:119493
   #map_position 11q22.1-11q22.3
CLASSIFICATION     #superfamily progesterone receptor, erbA transforming protein
                     homology
KEYWORDS           alternative splicing, DNA binding, glycoprotein; nucleus;
                     phosphoprotein; steroid hormone receptor; transcription
                     regulation; zinc finger
FEATURE
   1-933                 #product progesterone receptor form B #status predicted
                           #label MA1\
   165-933               #product progesterone receptor form A #status predicted
                           #label MA2\
   566-829               #domain erbA transforming protein homology #label ERBA\
   567-587               #region zinc finger CCCC motif\
   603-627               #region zinc finger CCCC motif\
   681-933               #domain steroid binding #status predicted #label SIB\
   41                    #binding_site phosphate (Thr) (covalent) #status
                           predicted\
   227,232,552,793       #binding_site phosphate (Ser) (covalent) #status
                           predicted\
   329,374,601           #binding_site phosphate (Tyr) (covalent) #status
                           predicted\
   664                   #binding_site carbohydrate (Asn) (covalent) #status
                           predicted
SUMMARY            #length 933  #molecular-weight [illegible]  #checksum 5865

  Query Match              18.9%,  Score 666,  DB 3,  Length 933;
  Best Local Similarity  33.4%;  Pred. No. 6.28e-97;
  Matches  110;  Conservative   89;  Mismatches 113;  Indels  17;  Gaps  17;

Db    566 iclicgdeasgchygvltcgsckvffkramegqhnylcagrndcivdkirrkncpacrlr 625
          .| :|:| ||| |||| .| .|| ||||...|:::|:|:: |:| :|| ||:| ||||
Qy     95 FCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLR 154

Db    626 kcrqagmvl-ggrkfk-kfnkvrvvraldavalpqplgvpnesqalsqrftfspgqdiql 683
          || : |||  |:|: :  :  ||  |: |    | :  | :: | : : |        |  |
Qy    155 KCYEVGMVKGGSRKERCGYRLVRRQRSADE-QL-HCAGKAKRSGGHAPRVR-ELLLDA-L 210

Db    684 lp-plinllmsiepdviyaghdntkpdtssslltslnqlgerqllsvvkwskslpgfrnl 742
           |  |:  |:  ||  :  :: :: | | :|:: ||  |:::|: :: |:' :||| :|
Qy    211 SPEQLVLTLLEAEPPHVLISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVEL 269

Db    743 hiddqitliqyswmslmvfglgwrsykhvsgqmlyfapdliln-eqrmk-essfyslcltm 801
           : ||: |::  || ::: || |||  | :| :: |||||:|: :  '   :: :   :
Qy    270 SLFDQVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDML 327

Db    802 wqipqefvklqvsqeeflcmkvllllntiplegl rsqtq-feemrs-sy-relikaiq- 857
                |  |  .  . |.||.| ..|||. ..  | : ||  :  |  : :  :  |:
Qy    328 LATTSRFRELKLQHKEYLCVKAMILLNS-SMYPLVTATQDADSSRKLAHLLNAVTDALVW 386

Db    858 lrqkgvvsssqrfyqltklldnlhdlvkq 886
           :  |: :|| |:  :|::||  |   |::
Qy    387 VIAKSGISSQQQSMRLANLL-MLLSHVRH 414


RESULT   15
ENTRY              A25923     #type complete
TITLE              progesterone receptor - rabbit
ORGANISM           #formal_name Oryctolagus cuniculus #common_name domestic
                     rabbit
DATE               16-Aug-1989 #sequence_revision 16 Aug 1989 #text_change
                     02-Jun-1996
ACCESSIONS         A25923
REFERENCE          A25923
   #authors        Loosfelt, H.; Atger, M.; Misrahi, M.; Guiochon-Mantel, A.;
                     Meriel, C.; Logeat, F.; Benarous, R.; Milgrom, E.
   #journal        Proc. Natl. Acad. Sci. U.S.A. (1986) 83:9045-9049
   #title          Cloning and sequence analysis of rabbit progesterone receptor
                     complementary DNA.
   #cross-references MUID:87067449
   #accession      A25923
      ##molecule_type mRNA
      ##residues      1-930 ##label LOO
CLASSIFICATION     #superfamily progesterone receptor, erbA transforming protein
                     homology
KEYWORDS           DNA binding, nucleus, steroid hormone receptor, zinc finger
FEATURE
   566-826               #domain erbA transforming protein homology #label ERBA\
   568-588               #region zinc finger\
   604-628               #region zinc finger
SUMMARY            #length 930  #molecular weight 98666  #checksum 2858
```

```
                  | ::   :   |  : ||: |: :|... | . |.   .|      ||.  |.||||.|
Qy     392 GISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNA 444


RESULT    11
ENTRY              S58087     #type complete
TITLE              estrogen receptor related protein - mouse
ORGANISM           #formal_name Mus musculus #common_name house mouse
DATE               13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change
                     26-Apr-1996
ACCESSIONS         S58087
REFERENCE          S58087
   #authors        Pettersson, K.; Svensson, K.; Mattsson, R.; Carlsson, B.;
                     Ohlsson, R.; Berkenstam, A.
   #submission     submitted to the EMBL Data Library, July 1995
   #description    Expression of a novel member of estrogen response
                     element-binding nuclear receptors is restricted to the
                     early stages of chorion formation during mouse
                     embryogenesis.
   #accession      S58087
      ##status           preliminary
      ##molecule_type mRNA
      ##residues         1-433 ##label PET
      ##cross-references EMBL:X89594
CLASSIFICATION   #superfamily erbA transforming protein homology
FEATURE
   101-352               #domain erbA transforming protein homology #label ERBA
SUMMARY          #length 433  #molecular-weight 48400  #checksum 1075

  Query Match              22.6%;  Score 799;  DB 14;  Length 433;
  Best Local Similarity   38.7%;  Pred. No. 4.93e-122;
  Matches   137;  Conservative   95;  Mismatches 110;  Indels   22;  Gaps   15;

Db     101 rlclvcgdiasgyhygvasceackaffkrtiqgniey ncpatnecei tkrrrksc qacrf 160
             ::| ||:| |||||||| |||:||||||||.|||. .| |||||:| | | ||||||||:
Qy      94 HFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRL 153

Db     161 mkclkvgmlkeqvrldrv--r-ggrqky-krrldsenspylnlpisppakkpltkivs-n 215
              ||  |||:| | | :|    |    ||:      :|        :  :   :|  .  |    :| :
Qy     154 RKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPE 213

Db     216 -l-lqv-eqdklyam-ppndipeqdikalttlceladrelvflinwakhipgfpsltlgd 271
             | | . | .      .  .         |   .      . .|  |||-|||  :|:|||  ||||    |:| |
Qy     214 QLVLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 273

Db     272 qmsllqsawmeililgivyrslpyddklayaedyimdeehsrlv-glldlyrailqlvrr 330
             |: ||:| |||:|..|.. ||.      || :| | ..| : .. | |.|... .|     |
Qy     274 QVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSR 333

Db     331 ykklkvekeefmilkalalansdsmy--i---enleavqklqdllh---ealq-dyels- 380
             :: ||:: |:: :||. | || |||    .      :: :. :||  ||.     .||        |
Qy     334 FRELKLQHKEYLCVKAMILLNS-SMYPLVTATQDADSSRKLAHLLNAVIDALVWVIAKSG 392

Db     381 -qrheeprragkllwtlpllrqtaakavqhfysvklqgkvpmhklflemleakv 433
                 :::: | ::||  |: :|::: |:::|: ::|      ||:  |:||||:| |
Qy     393 ISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNAHV 446


RESULT    12
ENTRY              A39596     #type complete
TITLE              progesterone receptor B form - mouse
ORGANISM           #formal_name Mus musculus #common_name house mouse
DATE               20-Mar-1992 #sequence_revision 20-Mar-1992 #text_change
                     03-Jun-1996
ACCESSIONS         A39596
REFERENCE          A39596
   #authors        Schott, D.R.; Shyamala, G.; Schneider, W.; Parry, G.
   #journal        Biochemistry (1991) 30:7014-7020
   #title          Molecular cloning, sequence analyses, and expression of
                     complementary DNA encoding murine progesterone receptor.
   #cross-references MUID:91299759
   #accession      A39596
      ##status           preliminary
```

```
      ##molecule_type mRNA
      ##residues         1-923 ##label SCH
      ##cross-references GB:J05333
CLASSIFICATION   #superfamily progesterone receptor, erbA transforming protein
                   homology
KEYWORDS         DNA binding; nucleus; steroid hormone receptor; zinc finger
FEATURE
   555-819               #domain erbA transforming protein homology #label ERBA\
   557-577               #region zinc finger\
   593-617               #region zinc finger
SUMMARY          #length 923  #molecular weight 99073  #checksum 6627

  Query Match              19.3%;  Score 681;  DB 7;  Length 923;
  Best Local Similarity   30.8%;  Pred. No. 9.61e-100;
  Matches   123;  Conservative  112;  Mismatches 138;  Indels   27;  Gaps   27;

Db     491 aapqtaaapaiyqplg-lnqlpqlg-y-qaavlkdslpqvyppylnylrpdseasqspqy 547
             ::||  ::   : : |: | : || :    :|   |   :|| |      |:   :   ::     ||
Qy      28 PTPGHLSPLVVHRQLSHLYAEPQKSPWCEARSLEHTLP-VNRETLK-RKVSGNRCASPVT 85

Db     548 gfdslpq-kiclicgdeasgchygvltcgsckvffkramegqhnylcagrndcivdkirr 606
             |   |   :   :| :|:| ||| |||| :| :|| ||||:::::::|:|:: |:| :|| ||
Qy      86 GPGSKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRR 145

Db     607 kncpacrlrkcceaqmvl-ggrkfk-kfnkvrvmrtldgvalpqsvglpnesqalsqrit 664
             |:| ||||||| : |||   |:|: :   :  ||   |: |     | :   | :: | : : |:
Qy     146 KSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQRSADE-QL-HCAGKAKRSGGHAPRVR 203

Db     665 fspnqeiqlvp plinllmsiepdviyaghdntkpdtsssllts lnqlgerqllsvvkws 723
                  .  | |  |:  |:  ||   :  :: ::  | | :|:: ||   |::::|: :: |:
Qy     204 -ELLLDA LSPEQLVLTLLEAEPPHVLISRPSA-PFTEASMMMSLTKLADKELVHMISWA 260

Db     724 kslpgfrnlhiddqitliqyswmslmvfqlqwrsykhvsqqmlyfapdliln-eq-r-mk 780
             | :||| :| : ||: |::   || ::: || |||   | :|  :: |||||:|:  : : :
Qy     261 KKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVE 318

Db     781 e-lsfyslcltmwqipqefvklqvtheeflcmkvlllnt i plegrsqsqfeemrss 837
               | :: : | :      |  | : | |:||:| ::|||: : ||        :::    : :
Qy     319 GILEIFDM-L-L-ATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAH 375

Db     838 yirelikaiq-lrqkgvvpssqrfyqltklldslhdlvkq 876
             :   :  |:   :  |: ::: |:  :|::||   |    |::
Qy     376 LLNAVIDALVWVIAKSGISSQQQSMRLANLL-MLLSHVRH 414


RESULT    13
ENTRY              I53280     #type complete
TITLE              progesterone receptor - rat
ORGANISM           #formal_name Rattus norvegicus #common_name Norway rat
DATE               02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change
                     02-Aug-1996
ACCESSIONS         I53280
REFERENCE          I53280
   #authors        Park-Sarge, O.K.; Mayo, K.E.
   #journal        Endocrinology (1994) 134:709-718
   #title          Regulation of the progesterone receptor gene by gonadotropins
                     and cyclic adenosine 3',5'-monophosphate in rat granulosa
                     cells.
   #cross-references MUID:94130817
   #accession      I53280
      ##status           preliminary, translated from GB/EMBL/DDBJ
      ##molecule_type mRNA
      ##residues         1-923 ##label RES
      ##cross-references GB:L16922; NID:g463282; CDS_PID:g463283
SUMMARY          #length 923  #molecular-weight 99407  #checksum 9843

  Query Match              19.0%;  Score 671;  DB 14;  Length 923;
  Best Local Similarity   32.6%;  Pred. No. 7.25e-98;
  Matches   108;  Conservative   94;  Mismatches 108;  Indels   21;  Gaps   21;

Db     556 iclicgdeasgchygvltcgsckvffkramegqhnylcagrndcivdkirrkncpacrlr 615
             :| :|:| ||| |||| :| :|| ||||:::|:::|:|:: |:| :|| |||:| |||||
Qy      95 FCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLR 154
```

```
      ##residues      1-701 ##label PIN
      ##cross-references EMBL:U47678
SUMMARY          #length 701  #molecular-weight 78430  #checksum 4951

  Query Match             41.7%;  Score 1473;  DB 16;  Length 701;
  Best Local Similarity   56.9%;  Pred. No. 4.00e-252;
  Matches  197;  Conservative   85;  Mismatches  49;  Indels  15;  Gaps  11;

Db    178 saketrycavcndyasgyhygvwscegckaffkrsiqghndymcpatnqctidknrrksc 237
          | :::::||||:|||||||||||||||||||||||||||||||:||||||||||||||||
Qy     89 SKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSC 148

Db    238 qacrlrkcyevgmmkggirkdrrggrmlkhkrqrddgegrgevgsagdmraanlwpsplm 297
           |::| ||||:|:| | |::| | |:::  |||    | :   :: : :|::     :|
Qy    149 QACRLRKCYEVGMVKCGSRRERCGYRLVR--RQRSADE-QLHCAGKA-KRSGG--HAP-R 201

Db    298 ikrskknslalsltadqmvsalldaepp-ilyseydptrpfseasmmglltnladrelvh 356
          : |       |  :|:::|:| :||:|||| :| |     |: ||:||||| ||:|||:||||
Qy    202 V-RELL--LD-ALSPEQLVLTLLEAEPPHVLISR--PSAPFTEASMMMSLTKLADKELVH 255

Db    357 minwakrvpgfvdltlhdqvhllecawleilmiglvwrsmehpvkllfapnllldrnqgk 416
          ||:||||::||||:|:| |||:|||  |:|:||:||:|||::|| ||:||||:|:|||::||
Qy    256 MISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGK 315

Db    417 cvegmveifdmllatssrfrmmnlqgeefvclksiillnsgvytflsstlksleekdhih 476
          ||||::||||||||||:||||    ||  |  |:|  |||||  |  :::|   :   :
Qy    316 CVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSRKLA 374

Db    477 rvldkitdtlihlmakagltlqqqhqrlaqllliLshirhmrnqgk 522
          ::|: :|:|: :.||.|.. |||  ||| ||..|||.||   | |:
Qy    375 HLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGM 420


RESULT    9
ENTRY             B29345     #type complete
TITLE             steroid hormone receptor ERR2 precursor - human
ORGANISM          #formal_name Homo sapiens #common_name man
DATE              15-Dec-1988 #sequence_revision 15-Dec-1988 #text_change
                    03-Jun-1996
ACCESSIONS        B29345
REFERENCE         A29345
   #authors       Giguere, V.; Yang, N.; Segui, P.; Evans, R.M.
   #journal       Nature (1988) 331:91-94
   #title         Identification of a new class of steroid hormone receptors.
   #cross-references MUID:88122546
   #accession     B29345
      ##molecule_type mRNA
      ##residues      1-433 ##label GIG
      ##cross-references GB:X51417; GB:Y00291
CLASSIFICATION    #superfamily erbA transforming protein homology
KEYWORDS          DNA binding; zinc finger
FEATURE
   101-352              #domain erbA transforming protein homology #label ERBA\
   103-123              #region zinc finger\
   139-163              #region zinc finger
SUMMARY          #length 433  #molecular-weight 48287  #checksum 1133

  Query Match             22.8%;  Score 807;  DB 13;  Length 433;
  Best Local Similarity   39.5%;  Pred. No. 1.49e-123;
  Matches  140;  Conservative   84;  Mismatches 109;  Indels  22;  Gaps  14;

Db    101 rlclvcgdiasgyhygvasceackaffkrtiqgnieyscpatneceitkrrrkscqacrf 160
          ::| ||:| ||||||||| |||:|||||||||.|||. .| |||||.| | | |||||||||:
Qy     94 HFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRL 153

Db    161 mkclkvgmlkegvrldrv--r-ggrqky-krrldsenspylslqisppakkpltkivs-- 214
            ||  ||:|| | | :|   |   ||:     :|      :  |    .|  .  |   .|
Qy    154 RKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPE 213

Db    215 yl-l-vaepdklyam-ppddvpegdikalttlcdladrelvfliswakhipgfsnltlgd 271
           | | : |::     : :     |   .     : :|  |||.||| :||||| |||| |.| |
Qy    214 QLVLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 273
```

```
Db    272 qmsllqsawmeilllgivyrslpyddklayaedyimdeehsrlv-gllelyrallqlvrr 330
          |: ||:| |||:|::|:: ||:     || :| | ::| : :: | |:||:: :|     |
Qy    274 QVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSR 333

Db    331 ykklkvek--efvtlkalalansdsmy--i---edleavqklqdllh---ealqdyels- 380
          :: |::: |:: :||: | || |||  :    :: :: :||  ||:   :||      |
Qy    334 FRELKLQHKEYLCVKAMILLNS-SMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKSG 392

Db    381 -qhleeprragkllltlpllrqtaakavqhfysvklggkvpmhklflemleakv 433
              :::: | ::||. |. .|... |...|. ..|    ||.  |.||||:| |
Qy    393 ISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNAHV 446


RESULT    10
ENTRY             A29345     #type complete
TITLE             steroid hormone receptor ERR1 precursor - human
ALTERNATE_NAMES   estrogen-related receptor
ORGANISM          #formal_name Homo sapiens #common_name man
DATE              15-Dec-1988 #sequence_revision 15-Dec-1988 #text_change
                    03-Jun-1996
ACCESSIONS        A29345; A49074
REFERENCE         A29345
   #authors       Giguere, V.; Yang, N.; Segui, P.; Evans, R.M.
   #journal       Nature (1988) 331:91-94
   #title         Identification of a new class of steroid hormone receptors.
   #cross-references MUID:88122546
   #accession     A29345
      ##molecule_type mRNA
      ##residues      1-521 ##label GIG
      ##cross-references EMBL:X51416; EMBL:Y00290
REFERENCE         A49074
   #authors       Wiley, S.R.; Kraus, R.J.; Zuo, F.; Murray, E.E.; Loritz, K.;
                    Mertz, J.E.
   #journal       Genes Dev. (1993) 7:2206-2219
   #title         SV40 early-to-late switch involves titration of cellular
                    transcriptional repressors.
   #accession     A49074
      ##status        preliminary
      ##molecule_type protein
      ##residues      166-169,'X',171-173 ##label WIL
CLASSIFICATION    #superfamily erbA transforming protein homology
KEYWORDS          DNA binding; nucleus; zinc finger
FEATURE
   174-434              #domain erbA transforming protein homology #label ERBA\
   176-196              #region zinc finger\
   212-236              #region zinc finger
SUMMARY          #length 521  #molecular-weight 55742  #checksum 5840

  Query Match             22.7%;  Score 800;  DB 13;  Length 521;
  Best Local Similarity   39.9%;  Pred. No. 3.19e-122;
  Matches  141;  Conservative   74;  Mismatches 128;  Indels  10;  Gaps   7;

Db    174 rlclvcgdvasgyhygvasceackaffkrtiqgsieyscpasneceitkrrrkacqacrf 233
          ::| ||:| ||||||||| |||:||||||||:|||  :| |||:|:| | | |||:||||||:
Qy     94 HFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRL 153

Db    234 tkclrvgmlkegvrldrvrggrqkykrrpevdplpfpg-p-fpagplavaggpktaapv 291
           ||  |||:| | | :|  | |   ::|:  : |  :| :  ::|  :          | :
Qy    154 RKCYEVGMVKCGSRRERC-GYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSP 212

Db    292 nalvshllvvepeklyampdpagpdghlpavatlcdlfdrelvvtiswakslpgfsslsl 351
          . ||  ||  ||  :  :: |::|     : : :|  | |:|:|  ||||| |||| |||
Qy    213 EQLVLTLLEAEPPHVL-ISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSL 271

Db    352 sdqmsvlqsvwmevlvlgvaqrslplqdelafaedlvld-eegaraaglgelgaallqlv 410
           ||. .|.| ||||||..|.  ||:       | || ||||| :||      |:|:   ||
Qy    272 FDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATT 331

Db    411 rrlqalrlereeyvllkalalansdsvhiedeprlwssceklllh--eal---leyeagra 465
          |.. |.|.: ||. :||: | ||      .      :     | ||| |  :|:    |    :::
Qy    332 SRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKS 391

Db    466 gpgggaerraqrllltlpllrqtadkvlahfyrvklegkvpmhklflemlea 518
```

```
             ||| ||||||||||:|||||||||:|| |::   | :  :   |::    :  :| :::|
Qy     421 EHLLNMKCKNVVPVYDLLLEMLNAHVLRGCKSSITGSECSPAEDSKSKEGSQNPQS 476


RESULT     6
ENTRY               A37197      #type complete
TITLE               estrogen receptor - rainbow trout
ORGANISM            #formal_name Oncorhynchus mykiss #common_name rainbow trout
DATE                31-Jan-1992 #sequence_revision 31-Jan-1992 #text_change
                      03-Jun-1996
ACCESSIONS          A37197; A40070
REFERENCE           A37197
   #authors         Pakdel, F.; Le Gac, F.; Le Goff, P.; Valotaire, Y.
   #journal         Mol. Cell. Endocrinol. (1990) 71:195-204
   #title           Full-length sequence and in vitro expression of rainbow trout
                      estrogen receptor cDNA
   #cross-references MUID:91006824
   #accession       A37197
      ##status            preliminary
      ##molecule_type mRNA
      ##residues          1-574 ##label PAK
      ##cross-references GB:M31559
REFERENCE           A40070
   #authors         Pakdel, F.; Le Guellec, C.; Vaillant, C.; Le Roux, M.G.;
                      Valotaire, Y.
   #journal         Mol. Endocrinol. (1989) 3:44-51
   #title           Identification and estrogen induction of two estrogen
                      receptors (ER) messenger ribonucleic acids in the rainbow
                      trout liver: sequence homology with other ERs.
   #cross-references MUID:89127284
   #accession       A40070
      ##status            preliminary
      ##molecule_type mRNA
      ##residues          150-574 ##label PA2
      ##cross-references GB:M31559
CLASSIFICATION      #superfamily estrogen receptor; erbA transforming protein
                      homology
KEYWORDS            DNA binding; nucleus; steroid hormone receptor; zinc finger
FEATURE
   145-421                #domain erbA transforming protein homology #label ERBA\
   147-167                #region zinc finger\
   183-207                #region zinc finger
SUMMARY             #length 574  #molecular-weight 62871  #checksum 1057

  Query Match              42.9%;  Score 1516;  DB 7;  Length 574;
  Best Local Similarity  56.3%;  Pred. No. 1.61e-260;
  Matches   225;  Conservative   78;  Mismatches  75;  Indels  22;  Gaps  17;

Db     143 etrycavcsdfasgyhygvwscegckaffkrsiqghndymcpatnqctmdrnrrkscqac 202
           ::::|||||||:|||||||||||||||||||||||||||||||:|||||||||:|:||||||||
Qy      92 DAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQAC 151

Db     203 rlrkcyevgmvkggirkdrgg-rvlrkdkrycgpagdrekptvtwstqgrpqdggrnsss 261
           |||||||||||| | |::| | |::|:: |    :| | |       :  :  |:  ||:   ::
Qy     152 RLRKCYEVGMVKCGSRRERCGYRLVRRQ-R---SA-D-EQLHCA-GKAKRS--GGH-AP 200

Db     262 slnggggwrgpritmppeqvlfllqgaeppalcsrqkvarpytevtmmtlltsmadkelv 320
             :        : :: | : || ||::: |   | ||    | |:|| :||  || :||||||
Qy     201 RVREL--LLDA-LS-PEQLVLTLLEAEPPHVLISRPS-A-PFTEASMMMSLTKLADKELV 254

Db     321 hmiawakkvpgfqelslhdqvqllesswlevlmiqliwrsihcpgklifaqdlildrseg 380
           |||:|||||:||| |||| |||:|||| |:|||||:||:|||| |||||||| ||:||| ||
Qy     255 HMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEG 314

Db     381 dcvegmaeifdmllatvsrfgmlklkpeefvclkaiillnpgafsfcsnsveslhnssav 440
            ||||: ||||||||| |||  |||   |::|:||:|||||:: ::: : : ::  :|  :
Qy     315 KCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSRKL 373

Db     441 esmldnitdalihhishsqasvqqqprrqaqlllllshirhmsnkgmehlysikcknkvp 500
             :|: :||||:  |: || | |||: | | ||:||||:|| |||||||| ::||||| ||
Qy     374 AHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVP 433

Db     501 lydlllemldqhrlqapqkvaqaeqteqpsttttttstgs 540
```

```
           :|||||||::| |::  | : :| :  :|:  : :  ||
Qy     434 VYDLLLEMLNAHVLRGC-KSSITGSEC-SPAEDSKSKEGS 471


RESULT     7
ENTRY               S58224      #type fragment
TITLE               oestrogen receptor - Atlantic salmon (fragment)
ORGANISM            #formal_name Salmo salar #common_name Atlantic salmon
DATE                13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change
                      26-Apr-1996
ACCESSIONS          S58224
REFERENCE           S58224
   #authors         Rogers, S.A.; Llewellyn, L.; Ramsurn, V.P.; Sweeney, G.E.;
                      Wigham, T.
   #submission      submitted to the EMBL Data Library, July 1995
   #accession       S58224
      ##status            preliminary
      ##molecule_type mRNA
      ##residues          1-535 ##label ROG
      ##cross-references EMBL:X89959
CLASSIFICATION      #superfamily erbA transforming protein homology
FEATURE
   103-382                #domain erbA transforming protein homology #label ERBA
SUMMARY             #length 535  #checksum 9953

  Query Match              42.6%;  Score 1503;  DB 7;  Length 535;
  Best Local Similarity  55.7%;  Pred. No. 5.57e-258;
  Matches   224;  Conservative   75;  Mismatches  80;  Indels  22;  Gaps  16;

Db     101 etrycavcsdfasgyhygvwscegckaffkrsiqghndympatnqctidknrrkscqac 160
           ::::|||||||:||||||| |||||||||||||||||||||||:||||||||:|:||||||||
Qy      92 DAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQAC 151

Db     161 rlrkcyevgmvkggirkdrgg-rvlrkdkrycqpagdrekpygdlehrtappqdggrnss 219
           |||||||||||| | |::| | |::|:: |    :| :  :    :    : |   : ||: :
Qy     152 RLRKCYEVGMVKCGSRRERCGYRLVRRQ-R---SADE-QLHCAG---K-AK-RSGGHAPR 201

Db     220 ssslsggggwrgpritmppeqvlfllqgaeppalcsrqkvarpytevtmmtlltsmadke 279
              |          : :: | : || ||::   |   | ||    | |:|| :||  || :||||
Qy     202 VRELL---L-DA-LS-PEQLVLTLLEAEPPHVLISRPS-A-PFTEASMMMSLTKLADKE 252

Db     280 lvhmiawakkvpgfqelslhdqvqllesswlevlmiqliwrsihcpgklifaqdlildrs 339
           |||||:||||:||| |||| |||:|||| |:|||||:||:||||  |||||||| ||:|||
Qy     253 LVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRD 312

Db     340 egdcvegmaeifdmllatvsrfrmlklkpeefvclkaiillnsgafsfcsnsveslhnss 399
           || ||||. ||||||||| |||| |||   |.:|:||:||||||: ::: : : ::  :|
Qy     313 EGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSR 371

Db     400 avesmldnitdalihhishsqasvqqqprrqvqlllllshirhmsnkgmehlysikcknk 459
            :  :|: :||||:  |: || | |||: |    ||:||||:|| |||||||| ::||||
Qy     372 KLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNV 431

Db     460 vplydlllemldghrlqspgkvaqaeqteqpsttttttstgs 501
           ||.:|||||||..| |::  | . .| .  .|.  .  .  ||
Qy     432 VPVYDLLLEMLNAHVLRGC-KSSITGSEC-SPAEDSKSKEGS 471


RESULT     8
ENTRY               S64737      #type complete
TITLE               80K estrogen receptor - human
ORGANISM            #formal_name Homo sapiens #common_name man
DATE                28-Oct-1996 #sequence_revision 28-Oct-1996 #text_change
                      28-Oct-1996
ACCESSIONS          S64737
REFERENCE           S64737
   #authors         Pink, J.J.; Wu, S.Q.; Wolf, D.M.; Bilimoria, M.M.; Jordan,
                      V.C.
   #journal         Nucleic Acids Res. (1996) 24:962-969
   #title           A novel 80 kDa human estrogen receptor containing a
                      duplication of exons 6 and 7.
   #accession       S64737
      ##status            preliminary
```

```
              proliferation and differentiation in target tissues.
COMMENT     In the absence of ligand, steroid hormone receptors are thought to
              be weakly associated with nuclear components; hormone binding
              greatly increases receptor affinity. The hormone-receptor complex
              appears to recognize discrete DNA sequences upstream of
              transcriptional start sites.
CLASSIFICATION   #superfamily estrogen receptor; erbA transforming protein
                   homology
KEYWORDS         DNA binding; nucleus; phosphoprotein, steroid binding,
                   steroid hormone receptor; transcription regulation; zinc
                   finger
FEATURE
   1-183                #domain amino-terminal #label NH2\
   184-275              #domain DNA binding #status predicted #label DNA\
   187-460              #domain erbA transforming protein homology #label ERBA\
   187-210              #region zinc finger CCCC motif\
   223-245              #region zinc finger CCCC motif\
   260-275              #region nuclear location signal\
   304-556              #domain steroid binding #status predicted #label STB\
   189,192,206,209      #binding_site zinc (Cys) #status predicted\
   225,231,241,244      #binding_site zinc (Cys) #status predicted\
   240,309              #binding_site phosphate (Ser) (covalent) #status
                          predicted
SUMMARY          #length 599  #molecular-weight 66954  #checksum 4749

  Query Match            47.3%;  Score 1669;  DB 3;  Length 599;
  Best Local Similarity  51.7%;  Pred. No. 1.92e-290;
  Matches  246;  Conservative  106;  Mismatches  104;  Indels  20;  Gaps  14;

Db    108 spsplmllhpppqlspflhphgqqvpyylenepsay-avrdtgppafyrsnsdnrrqngr 166
          : || :|   | :||| |  | |    | | : | :   |          :: :|: :|
Qy     20 TTSPNVLWPTPGHLSP-LVVHRQLSHLYAEPQKSPWCEAPSLEHTLPVNRETLKRKVSG- 77

Db    167 erlsssnekgnmimesaketrycavcndyasgyhygvwscegckaffkrsiqghndymcp 226
          :|  :|   |     | ::  :||||:||||||||||||||||||||||||||||||:||
Qy     78 NR-CASPVTGP---GSKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICP 133

Db    227 atnqctidknrrkscqacrlrkcyevgmmkggirkdrrggrmlkhkrqrddlegrnemga 286
          |:|||:||||||||||||||||||||||:| |  ::| | |::   |||   : :    |
Qy    134 AINQCIIDKNRRKSCQACRLRKCYEVGMVKCGSPRERCGYRLVP--RQRSA-DEQLH-CA 189

Db    287 sgdmraanlwpsplvikhtkknspalsltadqmvsalllaeppmlyseydpsrpfseasm 346
          :  :|::    :| | :    .  |||   :|:|  || |||| :     || ||:||||
Qy    190 GKAKRSGG--HAPRV-RELLLD--ALS--PEQLVLTLLEAEPPHVLIS-RPSAPFTEASM 241

Db    347 mglltnladrelvhminwakrvpgfgdlnlhdqvhllecawleilmiglvwrsmehpgkl 406
          |  ||:|||:||||||:|||::||| :|:| |||.|||  |:|:||:||:|||::|||||
Qy    242 MMSLTKLADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKL 301

Db    407 lfapnllldrnqgkcvegmveifdmllatssrfrmmnlqgeefvclksiillnsgvytfl 466
          :|||:|:|||::||||||::|||||||||:|||| ::||  |::|:|::|||||::|  :
Qy    302 IFAPDIVLDRDEGKCVEGILEIFDMLLATTSPFRELKLQHKEYLCVKAMILLNSSMYPLV 361

Db    467 sstlksleekdhihrvldkitdtlihlmakagltlqqqhrrlaqlllilshirhmsnkgm 526
           :|  : :   :: ::|: :|| |: ::||:|:: |||  |||:||::|||:|| |||||
Qy    362 IATQDA-DSSRKLAHLLNAVIDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGM 420

Db    527 ehlynmkcknvvplydlllemldahrlhapasrmgvppeepsqtqlattsstsahs 582
          ||| |||||||||:||||||||:|| |    |        |         |    |
Qy    421 EHLLNMKCKNVVPVYDLLLEMLNAHVLRGCKSSITGSECSPAEDSKSKEGSQNPQS 476


RESULT      5
ENTRY            QRRTE       #type complete
TITLE            estrogen receptor - rat
ORGANISM         #formal_name Rattus norvegicus #common_name Norway rat
DATE             30-Sep-1991 #sequence_revision 14-Jul-1994 #text_change
                   15-Jun-1996
ACCESSIONS       S07379; S16731
REFERENCE        S07379
   #authors      Koike, S., Sakai, M., Muramatsu, M.
   #journal      Nucleic Acids Res. (1987) 15:2499-2513
   #title        Molecular cloning and characterization of rat estrogen
                   receptor cDNA.
   #cross-references MUID:87174780
   #accession    S07379
      ##molecule_type mRNA
      ##residues       1-600 ##label KOI
      ##cross-references EMBL:Y00102
REFERENCE        S16731
   #authors      Maggi, A.M.A.
   #submission   submitted to the EMBL Data Library, June 1991
   #accession    S16731
      ##molecule_type mRNA
      ##residues       1-487,'T',489-600 ##label MAG
      ##cross-references EMBL:X61098
COMMENT     The steroid hormones and their receptors are involved in the
              regulation of eukaryotic gene expression and affect cellular
              proliferation and differentiation in target tissues.
COMMENT     In the absence of ligand, steroid hormone receptors are thought to
              be weakly associated with nuclear components, hormone binding
              greatly increases receptor affinity. The hormone-receptor complex
              appears to recognize discrete DNA sequences upstream of
              transcriptional start sites.
CLASSIFICATION   #superfamily estrogen receptor, erbA transforming protein
                   homology
KEYWORDS         DNA binding, nucleus, phosphoprotein, steroid binding;
                   steroid hormone receptor; transcription regulation; zinc
                   finger
FEATURE
   1-184                #domain amino-terminal #label NH2\
   185-276              #domain DNA binding #status predicted #label DNA\
   188-461              #domain erbA transforming protein homology #label ERBA\
   188-211              #region zinc finger CCCC motif\
   224-246              #region zinc finger CCCC motif\
   261-276              #region nuclear location signal\
   305-557              #domain steroid binding #status predicted #label STB\
   190,193,207,210      #binding_site zinc (Cys) #status predicted\
   226,232,242,245      #binding_site zinc (Cys) #status predicted\
   241,310              #binding_site phosphate (Ser) (covalent) #status
                          predicted
SUMMARY          #length 600  #molecular-weight 67029  #checksum 4599

  Query Match            47.1%;  Score 1663;  DB 3;  Length 600;
  Best Local Similarity  51.1%;  Pred. No. 2.67e-289;
  Matches  243;  Conservative  110;  Mismatches  103;  Indels  20;  Gaps  14;

Db    109 spsplmllhppphvspflhphghqvpyylenepsay-avrdtgppafyrsnsdnrrqnqr 167
          : || :|   | |:|| |  | :    | | : | :   |          :: :|: :|
Qy     20 TISPNVLWPIPGHLSP-LVVHRQLSHLYAEPQKSPWCEAPSLEHILPVNRETLKRKVSG- 77

Db    168 erlssssekgnmimesaketrycavcndyasgyhygvwscegckaffkrsiqghndymcp 227
          .|  .|.  |     | ...::||||:||||||||||||||||||||||||||||||:||
Qy     78 NR-CASPVTGP---GSKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICP 133

Db    228 atnqctidknrrkscqacrlrkcyevgmmkggirkdrrggrmlkhkrqrddlegrnemgt 287
          |:|||:||||||||||||||||||||||:| |  ::| | |::  :|  |: :     |
Qy    134 AINQCIIDKNRRKSCQACRLRKCYEVGMVKCGSPRERCGYRLVPRQRSADE-QLHCA-GK 191

Db    288 sgdmraanlwpsplvikhtkknspalsltadqmvsalldaeppllyseydpsrpfseasm 347
          .  .|..    .| | :    :  |||   :|:|  ||:|||| :     || ||:||||
Qy    192 A--KRSGG--HAPRV-RELLLD--ALS--PEQLVLILLEAEPPHVLIS-RPSAPFTEASM 241

Db    348 mglltnladrelvhminwakrvpgfgdlnlhdqvhllecawleilmiglvwrsmehpgkl 407
          |  ||:|||:||||||:|||::||| :|:| |||:|||  |:|:||:||:|||::|||||
Qy    242 MMSLIKLADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKL 301

Db    408 lfapnllldrnqgkcvegmveifdmllatssrfrmmnlqgeefvclksiillnsgvytfl 467
          .|||:|.|||  ||||||  |||||||||| |||| .||  | |.|.:||||||..| ..
Qy    302 IFAPDLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLV 361

Db    468 sstlksleekdhihrvldkindtlihlmakagltlqqqhrrlaqlllilshirhmsnkgm 527
           .|  . .   .. ..|. .|.| ..||.|:: |||  |||:||::|||:|| |||||
Qy    362 IATQDA-DSSRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGM 420

Db    528 ehlynmkcknvvplydlllemldahrlhapasrmgvppeepsqsqltttsstsahs 583
```

```
   #cross-references MUID:86122927
   #accession     A94284
      ##molecule_type mRNA
      ##residues        1-595 ##label GR1
      ##cross-references GB:M12674
REFERENCE         A93376
   #authors       Green, S., Walter, P., Kumar, V., Krust, A., Bornert, J.M.,
                    Argos, P.; Chambon, P.
   #journal       Nature (1986) 320:134-139
   #title         Human oestrogen receptor cDNA: sequence, expression and
                    homology to v-erb-A.
   #cross-references MUID:86146892
   #accession     A93376
      ##molecule_type mRNA
      ##residues        1-595 ##label GR2
      ##cross-references GB:X03635
REFERENCE         A43021
   #authors       Ponglikitmongkol, M.; Green, S.; Chambon, P.
   #journal       EMBO J. (1988) 7:3385-3388
   #title         Genomic organization of the human oestrogen receptor gene.
   #accession     A43021
      ##molecule_type DNA
      ##residues        142-161,205-225,244-264,356-374,'G',376,402-422,447-460,
                          'P',462-467,508-528 ##label PON
REFERENCE         S27140
   #authors       Keaveney, M.; Klug, J.; Gannon, F.
   #journal       DNA Seq. (1992) 2:347-358
   #title         Sequence analysis of the 5' flanking region of the human
                    estrogen receptor gene.
   #accession     S27143
      ##status          preliminary; translation not shown
      ##molecule_type DNA
      ##residues        1-115 ##label KEA
      ##cross-references EMBL:X62462
REFERENCE         S34000
   #authors       Pfeffer, U.; Fecarotta, E.; Castagnetta, L.; Vidali, G.
   #journal       Cancer Res. (1993) 53:741-743
   #title         Estrogen receptor variant messenger RNA lacking exon 4 in
                    estrogen-responsive human breast cancer cell lines.
   #accession     S34000
      ##status          preliminary
      ##molecule_type mRNA
      ##residues        216-254,367-399,'G',401-434 ##label PFE
      ##cross-references EMBL:X73067
      ##note            the authors translated the codon GGG for residue 400 as
                          Val
REFERENCE         A41925
   #authors       Dotzlaw, H.; Alkhalaf, M.; Murphy, L.C.
   #journal       Mol. Endocrinol. (1992) 6:773-785
   #title         Characterization of estrogen receptor variant mRNAs from
                    human breast cancers.
   #cross-references MUID:92293154
   #accession     A41925
      ##molecule_type mRNA
      ##residues        1-214,'ELPTLC' ##label DOT
      ##cross-references NCBIN:106580
      ##experimental_source clone 4; breast cancer
      ##note            sequence has been revised after extraction from NCBI
                          backbone
      ##note            the complete sequence of neither the nucleotide nor the
                          protein is shown in this paper
   #accession     B41925
      ##status          significant sequence differences
      ##molecule_type mRNA
      ##cross-references NCBIN:106597; GB:M69296
      ##experimental_source clone 24; breast cancer
COMMENT    The steroid hormones and their receptors are involved in the
             regulation of eukaryotic gene expression and affect cellular
             proliferation and differentiation in target tissues.
COMMENT    In the absence of ligand, steroid hormone receptors are thought to
             be weakly associated with nuclear components; hormone binding
             greatly increases receptor affinity. The hormone-receptor complex
             appears to recognize discrete DNA sequences upstream of
```

```
             transcriptional start sites.
GENETICS
   #gene          GDB:ESR
      ##cross-references GDB:119120
   #map_position 6q25.1-6q25.1
   #introns       151/2; 219/1; 254/1; 365/1; 412/2; 457/1; 518/2
CLASSIFICATION    #superfamily estrogen receptor; erbA transforming protein
                    homology
KEYWORDS          DNA binding; nucleus; phosphoprotein; steroid hormone
                    receptor; transcription regulation; zinc finger
FEATURE
   1-120                #domain amino-terminal #label NH2\
   121-299              #domain DNA binding #status predicted #label DNA\
   183-456              #domain erbA transforming protein homology #label ERBA\
   185-205              #region zinc finger CCCC motif\
   221-245              #region zinc finger CCCC motif\
   300-595              #domain steroid binding #status predicted #label STR\
   236,305              #binding_site phosphate (Ser) (covalent) #status
                          predicted
SUMMARY           #length 595  #molecular-weight 66258  #checksum 4317

  Query Match             47.7%;  Score 1683;  DB 3;  Length 595;
  Best Local Similarity   56.6%;  Pred. No. 3.47e-293;
  Matches  228;  Conservative   98;  Mismatches   60;  Indels   17;  Gaps   13;

Db     178 saketrycavcndyasgyhygvwscegckaffkrsiqghndymcpatnqctidknrrksc 237
           | .....||||.|||||||||||||||||||||||||||||:||||||||||||||||||
Qy      89 SKPDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSC 148

Db     238 qacrlrkcyevgmmkggirkdrrggrmlkhkrqrddgegrgevgsagdmraanlwpsplm 297
           |||||||||||||:| | |::| | |:::  |||    | :     :: : :|::     :|
Qy     149 QACRLRKCYEVGMVKCGSRRERCGYRLVR--RQRSADE-QLHCAGKA-KRSGG--HAP-R 201

Db     298 ikrskknslalsltadqmvsalldaepp-ilyseydptrpfseasmmglltnladrelvh 356
           : |     |  :|:::|:| :||:|||| :| |    |: ||:|||||  ||:||:||||
Qy     202 V-RELL--LD-ALSPEQLVLTLLEAEPPHVLISR--PSAPFTEASMMMSLTKLADKELVH 255

Db     357 minwakrvpgfvdltlhdqvhllecawleilmiglvwrsmehpgkllfapnllldrnqgk 416
           ||:|||::||||:|:| |||:|||  |:|:||:||:||||::|| ||:||||:|:||::||
Qy     256 MISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGK 315

Db     417 cvegmveifdmllatssrfrmmnlqgeefvclksiillnsgvytflsstlksleekdhih 476
           ||||..||||||||.|||| .:||  |::|:||::||||||::| ::::|  : :     :
Qy     316 CVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSRKLA 374

Db     477 rvldkitdtlihlmakagltlqqqhqrlaqlllilshirhmsnkgmehlysmkcknvvpl 536
           ..|. .||.|. ..||.|.: |||  ||| ||::|||:|| ||||||| :||||||||:
Qy     375 HLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPV 434

Db     537 ydlllemldahrlha-ptsrggasveetdqshlatagstsshs 578
           ||||||||:|| |::  :|  |:    :::|   :  ||  :::|
Qy     435 YDLLLEMLNAHVLRGCKSSITGSECSPAEDSK-SKEGSQNPQS 476


RESULT      4
ENTRY             QRMSF       #type complete
TITLE             estrogen receptor - mouse
ORGANISM          #formal_name Mus musculus #common_name house mouse
DATE              06-Mar-1992 #sequence_revision 14-Jul-1994 #text_change
                    15-Jun-1996
ACCESSIONS        A40061
REFERENCE         A40061
   #authors       White, R.; Lees, J.A.; Needham, M.; Ham, J.; Parker, M.
   #journal       Mol. Endocrinol. (1987) 1:735-744
   #title         Structural organization and expression of the mouse estrogen
                    receptor.
   #cross-references MUID:91042558
   #accession     A40061
      ##molecule_type mRNA
      ##residues        1-599 ##label WHI
      ##cross-references GB:M38651
COMMENT    The steroid hormones and their receptors are involved in the
             regulation of eukaryotic gene expression and affect cellular
  Query Match            48.8%;  Score 1723;  DB 3;  Length 596;
  Best Local Similarity  53.4%;  Pred. No. 5.03e-301;
  Matches  242;  Conservative  100;  Mismatches  94;  Indels  17;  Gaps  13;

Db     101 spvvflakipqlspflhhhqqvpyylenepsgytvreagppafyrsnsdnrrqsgrer 159
           || |:: :    :| |:: |: |       | |:: :         |              : :|: || :|
Qy      22 SPNVLWPTPGHLSPLVVHR-QLSHLYAEPQKSPWCEARSLEHTLPVNPEILKRKVSG-NR 79

Db     160 mssandkgppsmeshketrycavcsdyasgyhygvwscegckaffkrsiqghndymcpat 219
              ::    || :    | :::::|||||||||||||||||||||||||||||||||:||||
Qy      80 -CASPVTGP-G--SKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAT 135

Db     220 nqctidknrrksgqacrlrkcyevgmmkggirkdrrggrllkhkrqrddlegrgevgsae 279
           ||||||||||||||||||||||||||||:| | |::| | ||      ||:    ||  .
Qy     136 NQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVRR--RLSADEQLH-CAGKA 192

Db     280 irtasiwvnps-ksmklsp-lsltaeqlisalmeaeaplvysehdstkplseasmmtllt 339
            |:::    | |: : | : ||  :|||: :|:|||:| |      · |· |··||||| ||
Qy     193 KRSGG--HAPRVRELLLDA-LS--PEQLVLTLLEAEPPHVLISRPSA-PFTEASMMMSLT 246

Db     340 nladrelvhminwakrvpgfvdltlhdqvhllecawleilmvglIwrsvehpgklsfapn 399
           :||:||||||:'||::||||:|:| |||:|||  |:|:||:||:|||::||||| |||:
Qy     247 KLADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPD 306

Db     400 llldrnqgrcveglveifdmlvttatrfrmmrlrgeeficlksiillnsgvytflsstle 459
           |:|'|::|:|||:.|||'|'..|..|||  ..|.  |..'.|..||'||::| ..   :
Qy     307 LVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQD 366

Db     460 sledtdlihiildkiidtlvhfmaksglslqqqrrlaqlllilshirhmsnkgmehlys 519
           :      | |: |: : |:||  :||||:| |||  ||| ||::|||:|| |||||||| :
Qy     367 ADSSRKLAHL-LNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGMEHLLN 425

Db     520 mkcknvvplydlllemldahrihtpkdktttqe 552
            |:||||:|:| ||||:|| ::  |    |  |
Qy     426 MKCKNVVPVYDLLLEMLNAHVLRGCKSSITGSE 458


RESULT    2
ENTRY              QRCHE       #type complete
TITLE              estrogen receptor - chicken
ORGANISM           #formal_name Gallus gallus #common_name chicken
DATE               06-Mar-1992 #sequence_revision 14-Jul-1994 #text_change
                     15-Jun-1996
ACCESSIONS         A40914; S07192
REFERENCE          A40914
   #authors        Maxwell, B.L.; McDonnell, D.P.; Conneely, O.M.; Schulz, T.Z.;
                     Greene, G.L.; O'Malley, B.W.
   #journal        Mol. Endocrinol. (1987) 1:25-35
   #title          Structural organization and regulation of the chicken
                     estrogen receptor
   #cross-references MUID:88318621
   #accession      A40914
      ##status           not compared with conceptual translation
      ##molecule_type mRNA
      ##residues         1-589 ##label MAX
REFERENCE          S07192
   #authors        Krust, A.; Green, S.; Argos, P.; Kumar, V.; Walter, P.;
                     Bornert, J.M.; Chambon, P.
   #journal        EMBO J. (1986) 5:891-897
   #title          The chicken oestrogen receptor sequence: homology with v-erbA
                     and the human oestrogen and glucocorticoid receptors.
   #cross-references MUID:86247578
   #accession      S07192
      ##status           preliminary
      ##molecule_type mRNA
      ##residues         1-256,'E',258-589 ##label KRU
      ##cross-references EMBL:X03805
COMMENT     The steroid hormones and their receptors are involved in the
              regulation of eukaryotic gene expression and affect cellular
              proliferation and differentiation in target tissues.
COMMENT     In the absence of ligand, steroid hormone receptors are thought to
              be weakly associated with nuclear components; hormone binding
              greatly increases receptor affinity. The hormone-receptor complex
```

```
              appears to recognize discrete DNA sequences upstream of
              transcriptional start sites.
CLASSIFICATION  #superfamily estrogen receptor; erbA transforming protein
                  homology
KEYWORDS        DNA binding; nucleus; phosphoprotein; steroid binding;
                  steroid hormone receptor; transcription regulation; zinc
                  finger
FEATURE
   1-173              #domain amino-terminal #label NH2\
   174-265            #domain DNA binding #status predicted #label DNA\
   177-450            #domain erbA transforming protein homology #label ERBA\
   177-200            #region zinc finger C2C2 motif\
   214-238            #region zinc finger C2C2 motif\
   256-265            #region nuclear location signal\
   294-549            #domain steroid binding #status predicted #label STR\
   179,182,196,199    #binding_site zinc (Cys) #status predicted\
   215,221,231,234    #binding_site zinc (Cys) #status predicted\
   230,299            #binding_site phosphate (Ser) (covalent) #status
                        predicted
SUMMARY         #length 589  #molecular-weight 66773  #checksum 793

  Query Match            48.0%;  Score 1697;  DB 3;  Length 589;
  Best Local Similarity  54.3%;  Pred. No. 6.27e-296;
  Matches  242;  Conservative  99;  Mismatches  87;  Indels  18;  Gaps  11;

Db     100 spvvflqtapqlspflhhsqqvpyylenegssygmreaappatyrpsdnrrnsirerm 159
           || |:  |: :|||:: |: |      | | : ::        :    ::    : : :: :
Qy      22 SPNVLWPTPGHLSPLVVHR-QLSHLYAEPQKSPWCEARSLEHTL-PVNRETLKRKVSGNR 79

Db     160 sstnekgslsmestketrycavcndyasgyhygvwscegckaffkrsiqghndymcpatn 219
             ::    |: :   | ....||||.||||||||||||||||||||||||||||||:|'|:|
Qy      80 CASPVTGP-G--SKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATN 136

Db     220 qctidknrrkscqacrlrkcyevgmmkggirkdrrggrmlkhkrqrceqdsrngeasste 279
           ||||||||||||||||||||||||||:| | |::| | |:::: |   :||      | |  :
Qy     137 QCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSG 196

Db     280 lraptlwtsplvvkhnkknspalsltaeqmvsalleaepplvyseydpnrpfneasmmtl 339
             .||  ..   |..       .  |||  :||:| :|||||||  |        |: || |'|||
Qy     197 GHAPRVR-E-LLL---D--ALS-PEQLVLTLLEAEPPHVLIS-RPSAPFTEASMMMS 244

Db     340 ltnladrelvhminwakrvpgfvdltlhdqvhllecawleilmiglvwrsmehpgklifa 399
           ||:|||:||||||:|||::||||:|:| |||:|||  |:|:||:|||:|||::||||||:|
Qy     245 LTKLADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFA 304

Db     400 pnllldrnqgkcvegmveifdmllataarfrmmnlqgeefvclksiillnsgvytflss 459
           |:|:||||::|||||::||||||||:::|||  ::||  |::|:|::||||||::| ::::|
Qy     305 PDLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTAT 364

Db     460 lksleerdyihrvldkitdtlihlmaksglslqqqhrrlaqllllishirhmsnkgmehl 519
            .   |    | .|. .||.|. ..||||.| |||  ||| ||..||||:'| |||||| |
Qy     365 QDADSSRKLAH-LLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGMEHL 423

Db     520 ynmkcknvvplydlllemldahrlha 545
             |||||||||:|||||'|||:||' |::
Qy     424 LNMKCKNVVPVYDLLLEMLNAHVLRG 449


RESULT    3
ENTRY              QRHUE       #type complete
TITLE              estrogen receptor - human
ALTERNATE_NAMES    hER
ORGANISM           #formal_name Homo sapiens #common_name man
DATE               28-May-1986 #sequence_revision 28-May-1986 #text_change
                     15-Nov-1996
ACCESSIONS         A94284; A93376; A43011; S07143; S54000; A41925; B41925;
                     A03244; C41925; D41925
REFERENCE          A94284
   #authors        Greene, G.L.; Gilna, P.; Waterfield, M.; Baker, A.; Hort, Y.;
                     Shine, J.
   #journal        Science (1986) 231:1150-1154
   #title          Sequence and expression of human estrogen receptor
                     complementary DNA.
```

```
     ********************************************************************
          __    ___   ______  _____   _____  _____   _    _
         | |\  /|  | |  ___ | |  ___|| |  __ || |  __| | |  | |
         | | \/ |  | | |___|| | |___ | | |__|| | |     | |__| |
         | |    |  | |  ____| |___  || |  __ | | |     |  __  |
         | |    |  | | |          | || | |\ \| | |     | |  | |
         | |    |  | | |       ___| || | | \ \ | |___  | |  | |
         |_|    |__| |_|      |_____||_|  \_\|_____| |_|  |_|  (TM)

     ********************************************************************

          Release 2.1D John F. Collins, Biocomputing Research Unit.
          Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
                 Distribution rights by IntelliGenetics, Inc.

MPsrch_pp   protein - protein database search, using Smith-Waterman algorithm

Run on:          Thu Dec 18 08:42:55 1997,  MasPar time 17.36 Seconds
                                            793.885 Million cell updates/sec
ular output not generated.

Title:           >US-08-826-361A-5
Description:     (1-477) from US08826361A.pep
Perfect Score:   3532
Sequence:        1 MNYSIPSNVTNLEGGPGRQT          FCSPAEDSKSKEGSQNPQSQ 477

Scoring table:   PAM 150
                 Gap 11

Searched:        91006 seqs, 28888923 residues

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database:        pir51
                   1:ann1 2:ann2 3:ann3 4:ann4 5:unann1 6:unann2 7:unann3
                   8:unann4 9:unann5 10:unann6 11:unann7 12:unann8
                   13:unann9 14:unann10 15:unenc 16:unrev

Statistics:      Mean 48.118;  Variance 103.762;  scale 0.464

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution

                                  SUMMARIES
                   %
ult              Query
 No     Score    Match Length DB  ID              Description            Pred. No.
-----------------------------------------------------------------------------------
    1    1723    48.8    586   3  QRXLE           estrogen receptor      5.04e-301
    2    1697    48.0    599   3  QRCHE           estrogen receptor -    6.27e-296
    3    1683    47.7    595   3  QRHUE           estrogen receptor -    3.47e-293
    4    1669    47.3    599   3  QRMSE           estrogen receptor -    1.32e-290
    5    1663    47.1    600   3  QRRTE           estrogen receptor -    2.87e-289
    6    1516    42.9    574   7  A37197          estrogen receptor -    1.61e-260
    7    1503    42.6    535   7  S58224          oestrogen receptor -   5.57e-258
    8    1473    41.7    701  16  S64737          80K estrogen recepto   4.00e-252
    9     807    22.8    433  13  B29345          steroid hormone rece   1.49e-123
   10     803    22.7    521  13  A29345          steroid hormone rece   1.15e-122
   11     799    22.6    433  14  S58087          estrogen receptor re   4.96e-122
   12     681    19.3    923   7  A33596          progesterone recepto   9.61e-100
   13     671    19.0    933  14  I52290          progesterone recepto   7.25e-98
   14     666    18.9    933   3  QRHUP           progesterone recepto   6.26e-97
   15     662    18.7    931   7  A25719          progesterone recepto   3.54e-96
   16     614    17.4    742   3  QRHUGB          glucocorticoid recep   3.32e-87
   17     614    17.4    777   3  QRHUGA          glucocorticoid recep   3.32e-87
   18     608    17.2    776   7  S44047          glucocorticoid recep   4.35e-86
   19     597    16.9    786   7  A35466          progesterone recepto   4.85e-84
   20     594    16.8    771   7  A54273          glucocorticoid recep   1.75e-82
   21     586    16.6    783   3  A25691          glucocorticoid recep   5.37e-82
```

```
   22     579    16.4    123   7  S26595          estrogen receptor      1.07e-80
   23     577    16.3    795   3  QRRTG           glucocorticoid recep   2.51e-80
   24     571    16.2    899  14  A25895          androgen receptor -    3.25e-79
   25     570    16.1    902  14  B40444          androgen receptor -    4.38e-79
   26     567    16.1    919  13  A39248          androgen receptor -    1.79e-78
   27     564    16.0    910  13  A44221          androgen receptor A    1.42e-78
   28     560    15.9    911  13  B34721          androgen receptor B    3.53e-77
   29     562    15.9    981  14  A41401          mineralocorticoid re   1.51e-77
   30     543    15.4    344  13  I51330          androgen receptor -    4.87e-74
   31     544    15.4    758  16  S60586          glucocorticoid recep   3.18e-74
   32     533    15.1    533  13  S37781          retinoid X receptor    3.39e-72
   33     529    15.0    410  14  S26669          retinoic acid recept   1.85e-71
   34     529    15.0    446  14  A34418          H-2 region II bindin   1.85e-71
   35     529    15.0    448  14  D41727          retinoid X receptor    1.85e-71
   36     529    15.0    451  14  A41651          retinoic acid recept   1.85e-71
   37     523    14.8    984  13  A29513          mineralocorticoid re   2.34e-70
   38     514    14.6    466  14  C41727          retinoid X receptor    1.05e-68
   39     514    14.6    467  14  S26668          retinoic acid recept   1.05e-68
   40     514    14.6    467  14  A47278          retinoid X receptor    1.05e-68
   41     513    14.5    462  13  S09542          retinoic acid recept   1.61e-68
   42     500    14.2    385  14  I53158          orphan nuclear recep   3.87e-66
   43     502    14.2    563   7  JQ0523          nerve growth factor-   1.66e-66
   44     501    14.2    598   7  A37251          probable nuclear hor   2.54e-66
   45     499    14.1    488  13  C41977          retinoid receptor al   5.89e-66

                                  ALIGNMENTS

RESULT     1
ENTRY            QRXLE      #type complete
TITLE            estrogen receptor - African clawed frog
ORGANISM         #formal_name Xenopus laevis #common_name African clawed frog
DATE             28-Feb-1992 #sequence_revision 14-Jul-1994 #text_change
                   15-Jun-1996
ACCESSIONS       A40907
REFERENCE        A40907
   #authors      Weiler, I.J.; Lew, D.; Shapiro, D.J.
   #journal      Mol. Endocrinol. (1987) 1:355-362
   #title        The Xenopus laevis estrogen receptor: sequence homology with
                   human and avian receptors and identification of multiple
                   estrogen receptor messenger ribonucleic acids.
   #cross-references MUID:90331927
   #accession    A40907
      ##molecule_type mRNA
      ##residues      1-586 ##label WEI
      ##cross-references GB:L20735
COMMENT    The steroid hormones and their receptors are involved in the
             regulation of eukaryotic gene expression and affect cellular
             proliferation and differentiation in target tissues.
COMMENT    In the absence of ligand, steroid hormone receptors are thought to
             be weakly associated with nuclear components; hormone binding
             greatly increases receptor affinity. The hormone receptor complex
             appears to recognize discrete DNA sequences upstream of
             transcriptional start sites.
CLASSIFICATION   #superfamily estrogen receptor; erbA transforming protein
                   homology
KEYWORDS         DNA binding; nucleus; phosphoprotein; steroid binding;
                   steroid hormone receptor; transcription regulation; zinc
                   finger
FEATURE
   1-174                #domain amino-terminal #label NBD\
   175-266              #domain DNA binding #status predicted #label DNA\
   179-249              #domain erbA transforming protein homology #label ERBA\
   178-201              #region zinc finger CCCC motif\
   214-237              #region zinc finger CCCC motif\
   251-266              #region nuclear location signal\
   272-544              #domain steroid binding #status predicted #label STB\
   180,183,197,200      #binding_site zinc (Cys) #status predicted\
   216,222,234,237      #binding_site zinc (Cys) #status predicted\
   231                  #binding_site phosphate (Ser) (covalent) #status
                          predicted
SUMMARY          #length 586  #molecular-weight 66080  #checksum 9991
```

```
Search completed: Thu Dec 18 08:46:00 1997
Job time : 62 secs.
```

```
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS
CC   -!  ALTERNATIVE PRODUCTS: ALPHA AND BETA FORMS OF HGR DIFFER ONLY IN
CC       THEIR CARBOXYL-TERMINAL SEQUENCES  THEY ARE GENERATED BY
CC       ALTERNATIVE SPLICING. ALPHA-HGR IS THE PREDOMINANT PHYSIOLOGICAL
CC       FORM OF HGR  THE IN VITRO TRANSLATION PRODUCT OF HGR BETA DOES NOT
CC       BIND STEROID ITS FUNCTION IS UNCLEAR; HOWEVER, IT IS POSSIBLE THAT
CC       VARIANT RECEPTORS PERFORM TISSUE-SPECIFIC FUNCTIONS
CC   -!- DISEASE: DEFECTS IN GRL ARE A CAUSE OF FAMILIAL GLUCOCORTICOID
CC       RESISTANCE (CORTISOL RESISTANCE); A HYPERTENSIVE, HYPERANDROGENIC
CC       DISORDER CHARACTERIZED BY INCREASED SERUM CORTISOL CONCENTRATIONS.
DR   EMBL; M10901; G183033; -.
DR   EMBL; M11050; G183035; -
DR   EMBL; U01351; G458657; -.
DR   EMBL, X03225, G31680, -.
DR   EMBL, X03348, G31682, -
DR   EMBL; M69104; G553322; -
 R   EMBL; M73816; G553323, -.
     EMBL; S68378; G239758; -
     PIR; A03246; QRHUGA.
DR   PIR; A03247; QRHUGB
DR   HSSP; P06536; 1GDC
DR   MIM; 138040; -.
DR   PROSITE, PS00031; STEROID_FINGER
KW   RECEPTOR, TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING; DISEASE MUTATION; POLYMORPHISM;
KW   ALTERNATIVE SPLICING
FT   DOMAIN       145    278       IMMUNOGENIC.
FT   DNA_BIND     421    486       C4-TYPE ZINC FINGERS (TWO)
FT   ZN_FING      421    441       C4-TYPE.
FT   ZN_FING      457    481       C4-TYPE.
FT   DOMAIN       487    527       HINGE.
FT   DOMAIN       528    777       STEROID-BINDING.
FT   VARIANT      363    363       N -> S
FT   VARIANT      641    641       D -> V (IN GLUCOCORTICOID RESISTANCE).
FT   VARIANT      729    729       V -> I (IN GLUCOCORTICOID RESISTANCE)
FT   VARIANT      753    753       L -> F (IN GLUCOCORTICOID RESISTANCE).
FT   VARSPLIC     728    777       VVENLLNYCFQTFLDKTMSIEFPEMLAEIITNQIPKYSNGN
FT                                 IKKLLFHQK -> ENVMWLKPESTSHTLI (IN BETA
FT                                 FORM).
FT   CONFLICT     451    451       G -> GR (IN REF  2)
SQ   SEQUENCE   777 AA;  85659 MW;  D93D1F01 CRC32;

  Query Match             17.4%;  Score 614;  DB 4;  Length 777;
  Best Local Similarity   33.4%;  Pred. No. 1.07e-111;
  Matches   97;  Conservative   84;  Mismatches   93;  Indels   16;  Gaps   15;

       397 smrpdvssppsssstattgp-p-pklclvcsdeasgchygvltcgsckvffkravegqhn 454
           :::  ||:    :|: :  |:     : :| |||| ||| |||| :| :|| ||||::|:::
Qy      70 TLKRKVSGNPCASPVTGPGSKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHND 129

Db     455 ylcagrndciidkirrkncpacryrkclqagmnl-earktk-k-kikgiqqattg-v-sq 509
           |:|:: |:| ||| |||:| ||| ||| : ||      :|: :    ::   |:::    :
Qy     130 YICPATNQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCA 189

Db     510 etsenpgnkt-ivpatlpq-ltpt-lv-sllevieepevlyagydssvpdstwrimttlnm 565
             :  :|  :  |    | : |:| || :|||    | || :    ::  :::   :| :| :
Qy     190 GKAKRSGGHAPRVRELLLDALSPEQLVLTLLEAEPPHVLISRPSAPFTEAS--MMMSLTK 247

Db     566 lggrqviaavkwakaipgfrnlhlddqmtllqyswmflmafalgwrsyrqssanllcfap 625
           |: ::::  : ||| |||| :| | ||: ||:  ||  ::  :| |||   :::  :|: |||
Qy     248 LADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPG-KLI-FAP 305

Db     626 dliin-eqrmtlpcmydqckhmlyvsselhrlqvsyeeylcmktllllss 674
           ||::: :: : :  : :    :|  :| :: | :   ||||:|:::||:|
Qy     306 DLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNS 355


RESULT   15
ID   GCR_XENLA      STANDARD;      PRT;   776 AA.
```

```
AC   P49844;
DT   01-OCT-1996 (REL. 34, CREATED)
DT   01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   GLUCOCORTICOID RECEPTOR (GR).
OS   XENOPUS LAEVIS (AFRICAN CLAWED FROG).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AMPHIBIA; ANURA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=EMBRYO;
RX   MEDLINE; 94289478.
RA   GAO X., KALKHOVEN E., PETERSON-MADURO J., VAN DER BURG B.,
RA   DESTREE O.H.J.;
RL   BIOCHIM. BIOPHYS. ACTA 1218:194-198(1994).
RN   [2]
RP   SEQUENCE OF 452-776 FROM N.A.
RC   TISSUE=LIVER;
RA   PICARD D.;
RL   SUBMITTED (FEB-1994) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES. BINDS TO THE
CC       GRE TARGET SITE.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; X72211; G444043; -.
DR   EMBL; X77764; G456374; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER, STEROID-BINDING.
FT   DOMAIN         1    419       MODULATING.
FT   DNA_BIND     420    485       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING      420    440       C4-TYPE.
FT   ZN_FING      456    480       C4-TYPE.
FT   DOMAIN       486    776       STEROID-BINDING.
FT   CONFLICT     502    502       S -> F (IN REF. 2).
FT   CONFLICT     504    504       T -> A (IN REF  2).
FT   CONFLICT     506    506       T -> A (IN REF. 2).
FT   CONFLICT     511    511       P -> N (IN REF. 2).
FT   CONFLICT     519    519       L -> M (IN REF. 2).
FT   CONFLICT     544    544       Y -> F (IN REF. 2).
FT   CONFLICT     551    551       I -> M (IN REF. 2).
FT   CONFLICT     654    654       S -> R (IN REF. 2).
FT   CONFLICT     740    740       L -> M (IN REF. 2).
FT   CONFLICT     768    768       L -> I (IN REF. 2).
SQ   SEQUENCE   776 AA;  84977 MW;  896C496D CRC32;

  Query Match             17.2%;  Score 608;  DB 4;  Length 776;
  Best Local Similarity   34.9%;  Pred. No. 2.90e-110;
  Matches   99;  Conservative   74;  Mismatches   96;  Indels   15;  Gaps   15;

Db     398 irsdaspspstsststgpppklclvcsdeasgchygvltcgsckvffkravegqhnylca 457
           : ::  :|| |:: |     : :| |||| ||| |||| :| :|| ||||:::|:::|:|:
Qy      75 VSGNPCASPVTGPGSK-RDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICP 133

Db     458 grndciidkirrkncpacryrkclqagmnl-earktk-k-kikgiqqsttatar-es-pe 512
           : |:| ||| |||:| ||| ||| : ||      :|: :    ::   |:|:      :  : :
Qy     134 ATNQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAK 193

Db     513 tsmtrtl-vpasv-aqltpt-li-sllevieepevlysgydssipdttrrlmsslnmlggr 568
            |  ::  |   :    |:|  |: :|||    | || |   ::: :::   :|| | :|: :
Qy     194 RSGGHAPRVRELLLDALSPEQLVLTLLEAEPPHVLISRPSAPFTEASM-MMS-LTKLADK 251

Db     569 qvvsavrwakaipgfrnlhlddqmtllqyswmflmvfalgwrsykqtnqsllyfapdlvi 628
           ..|  . ||| |||| :| | ||: ||:  || ::: :| |||   :   |   | ||||||:
Qy     252 ELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHP-GK-LIFAPDLVL 309

Db     629 t-edrmhlpfmqercqemlkiaqemsslqisydeylcmkvlllm 671
             :: : :   : |   : :|  ::      | :    ||||:| ::|:
Qy     310 DRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILL 353
Qy     155 KCYEVGMVKGGSRRERCGYRLVRRQRSADE-QL-HCAGKAKRSGGHAPRVR-ELLLDA-L 210

Db     684 ip-plinllmsiepdviyaghdntkpdtsssllltslnqlgerqllsvvkwskslpgfrnl 742
             |  |:  |:  ||   :   :: :: | | :|  ||  |    .|. :: |:| :||| :|
Qy     211 SPEQLVLTLLEAEPPHVLISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVEL 269

Db     743 hiddqitllqyswmslmvfglgwrsykhvsgqmlyfapdlilneqrmk-essfyslcltm 801
             : ||: |::  || ::: || |||  | :| :: ||||:|: :  |   :: :   :
Qy     270 SLFDQVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDML 327

Db     802 wqipqefvklqvsqeeflcmkvllllntipleglrsqtq-feemrs-syi-relikaig- 857
                  |   | : : |:||:| ::|||: :: | : ||   .  |  . :   : |:
Qy     328 LATTSRFRELKLQHKEYLCVKAMILLNS-SMYPLVTATQDADSSRKLAHLLNAVTDALVW 386

Db     858 lrqkgvvsssqrfyqltklldnlhdlvkq 886
            :  |: :|| |:  :|::||  |   |::
Qy     387 VIAKSGISSQQQSMPLANLL-MLLSHVRH 414
```

```
RESULT   13
ID   PRGR_RABIT     STANDARD;      PRT;   930 AA.
AC   P06186;
DT   01-JAN-1988 (REL. 06, CREATED)
DT   01-JAN-1988 (REL. 06, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   PROGESTERONE RECEPTOR (PR).
GN   PGR.
OS   ORYCTOLAGUS CUNICULUS (RABBIT).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; LAGOMORPHA.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 87067449.
RA   LOOSFELT H., ATGER M., MISRAHI M., GUIOCHON-MANTEL A., MERIEL C.,
RA   LOGEAT F., BENAROUS R., MILGROM E.;
RL   PROC. NATL. ACAD. SCI. U.S.A. 83:9045-9049(1986).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS. A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS
DR   EMBL; M14547; G165632; -.
DR   PIR; A25923; A25923.
DR   HSSP; P06536; 1GDC.
     PROSITE; PS00031; STEROID_FINGER.
     TRANSFAC; T00697; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
FT   DOMAIN        1    565       MODULATING, PRO-RICH.
FT   DNA_BIND    568    633       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     568    588       C4-TYPE.
FT   ZN_FING     604    628       C4-TYPE.
FT   DOMAIN      678    930       STEROID-BINDING.
FT   DOMAIN      184    188       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
SQ   SEQUENCE   930 AA;  98666 MW;  EE133E86 CRC32;

  Query Match             18.7%;  Score 662;  DB 7;  Length 930;
  Best Local Similarity   32.5%;  Pred. No. 3.48e-123;
  Matches  106;  Conservative   91;  Mismatches 114;  Indels  15;  Gaps  15;

Db     567 iclicgdeasgchygvltcgsckvffkramegqhnylcagrndcivdkirrkncpacrlr 626
           :| :|:| ||| |||| :| :|| |||:::|...|.|.: |:| :|| |||:| |||||
Qy      95 FCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCIIDKNRRKSCQACRLR 154

Db     627 kccqagmvl-ggrkfk-kfnkvrvmraldavalpqpvgipnesq-ritfspsqeiq-lip 682
           || : |||  |:|: .  .  ||  |. |   | :  | : . |              . | |
Qy     155 KCYEVGMVKGGSRRERCGYRLVRRQRSADE-QL-HCAGKAKRSGGHAPRVRELLLDALSP 211

Db     683 -plinllmsiepdviyaghdntkpdtsssllltslnqlgerqllsvvkwskslpgfrnlhi 741
              |:  |:  ||   :   :: :: | | :|.. ||  |....|. .. |:| :||| :| :
```

```
Qy     213 EQLVLTLLEAEPPHVLISRPSA-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSL 271

Db     742 ddqitllqyswmslmvfglgwrsykhvsgqmlyfapdlilneqrmk-essfyslcltmwq 800
             ||. |.. || ... || |||  | .| .. |||||.|. .  |   .. :  :
Qy     272 FDQVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVEGILEIFDMLLA 329

Db     801 ipqefvklqvsqeeflcmkvllllnt-ipleglrsqtqfeemrssyirelikaiglrq 857
                 |  | . . |.||.| .:|||. . ||    :::  : :  :  : |: :
Qy     330 TISRPFELKLQHKEYLCVKAMILLNSSMYPLVIAIQDADSSRKLAHLLNAVIDALVWVIA 389

Db     858 kgvvsssqrfyqltklldnlhdlvkq 883
           |: :|| |:  :|::||  |   |::
Qy     390 KSGISSQQQSMPLANLL-MLLSHVRH 414

RESULT   14
ID   GCR_HUMAN      STANDARD;      PRT;   777 AA.
AC   P04150; P04151;
DT   01-NOV-1986 (REL. 03, CREATED)
DT   01-NOV-1986 (REL. 03, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   GLUCOCORTICOID RECEPTOR (GR).
GN   GRL.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 86092206.
RA   HOLLENBERG S.M., WEINBERGER C., ONG E.S., CERELLI G., ORO A.,
RA   LEBO R., THOMPSON E.B., ROSENFELD M.G., EVANS R.M.;
RL   NATURE 318:635-641(1985).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   MUNROE D.G., PANG J., TAYLOR G.R., LAU C., PLANTE R.K., ZHOU L.;
RL   SUBMITTED (SEP-1993) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [3]
RP   SEQUENCE OF 1-394 FROM N.A.
RX   MEDLINE; 91224961.
RA   LECLERC S., XIE B.X., ROY R., GOVINDAN M.V.;
RL   J. BIOL. CHEM. 266:8711-8719(1991).
RN   [4]
RP   DOMAINS.
RX   MEDLINE; 86092211.
RA   WEINBERGER C., HOLLENBERG S.M., ROSENFELD M.G., EVANS R.M.;
RL   NATURE 318:670-672(1985).
RN   [5]
RP   VARIANT SER-363.
RX   MEDLINE; 93187003.
RA   KARL M., LAMBERTS S.W.J., DETERA-WADLEIGH S.D., ENCIO I.J.,
RA   STRATAKIS C.A., HURLEY D.M., ACCILI D., CHROUSOS G.P.;
RL   J. CLIN. ENDOCRINOL. METAB. 76:683-689(1993).
RN   [6]
RP   VARIANT VAL-641.
RX   MEDLINE; 91123468.
RA   HURLEY D.M., ACCILI D., STRATAKIS C.A., KARL M., VAMVAKOPOULOS N.,
RA   RORER E., CONSTANTINE K., TAYLOR S.I., CHROUSOS G.P.;
RL   J. CLIN. INVEST. 87:680-686(1991).
RN   [7]
RP   VARIANT ILE-729.
RX   MEDLINE; 93253031.
RA   MALCHOFF D.M., BRUFSKY A., REARDON G., MCDERMOTT P., JAVIER E.C.,
RA   BERGH C.H., ROWE D., MALCHOFF C.D.;
RL   J. CLIN. INVEST. 91:1918-1925(1993).
RN   [8]
RP   VARIANT PHE-753.
RX   MEDLINE; 93302771.
RA   ASHRAF J., THOMPSON E.B.;
RL   MOL. ENDOCRINOL. 7:631-642(1993).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES. BINDS TO THE
CC       GRE TARGET SITE.
```

```
Qy     392 GISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNA 444

RESULT   11
ID    PRGR_MOUSE     STANDARD;      PRT;   923 AA.
AC    Q00175;
DT    01-DEC-1992 (REL. 24, CREATED)
DT    01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)
DT    01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE    PROGESTERONE RECEPTOR (PR)
GN    PGR OR PR.
OS    MUS MUSCULUS (MOUSE).
OC    EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC    EUTHERIA; RODENTIA
RN    [1]
RP    SEQUENCE FROM N.A
RX    MEDLINE; 91299759.
RA    SCHOTT D.R., SHYAMALA G., SCHNEIDER W., PARRY G.;
RL    BIOCHEMISTRY 30:7014-7020(1991)
CC    -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
          THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
          PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC    -!- SUBCELLULAR LOCATION: NUCLEAR
CC    -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC        A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC    -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC        NUCLEAR HORMONE RECEPTORS.
DR    EMBL; M68915; G200472; -.
DR    PIR; A39596; A39596.
DR    HSSP; P06536; 1GDC.
DR    PROSITE; PS00031; STEROID_FINGER.
KW    RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW    ZINC-FINGER; STEROID-BINDING.
FT    DOMAIN        1    556       MODULATING, PRO-RICH.
FT    DNA_BIND    557    622       C4-TYPE ZINC FINGERS (TWO).
FT    ZN_FING     557    577       C4-TYPE.
FT    ZN_FING     593    617       C4-TYPE.
FT    DOMAIN      671    923       STEROID-BINDING.
FT    DOMAIN      184    188       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
SQ    SEQUENCE   923 AA;  99073 MW;  EEDEF41A CRC32;

  Query Match            19.3%;  Score 681;  DB 7;  Length 923;
  Best Local Similarity  30.8%;  Pred. No. 9.45e-128;
  Matches   123;  Conservative  112;  Mismatches 128;  Indels  27;  Gaps  27;

Db     491 aapgtaaapaiyqplg-lnglpqlg-y-qaavlkdslpqvyppylnylrpdseasqspqy 547
           ::||  ::  : : |: | : || :    :|  |  :||  |    |:  :  :.   ||
Qy      28 PTPGHLSPLVVHRQLSHLYAEPQKSPWCEARSLEHTLP-VNRETLK-RKVSGNRCASPVT 85

Db     548 qfdslpq-kicliqqdeasqchygvltcgsckvffkrameqqhnylcagrndcivdkirr 606
           |  |  :  :| :|:| ||| |||| :| :|| ||||::.|...|.|.. |.| .|| ||
Qy      86 GPGSKRDAHRCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRR 145

Db     607 kncpacrlrkccqagmvl-ggrkfk-kfnkvrvmrtldgvalpqsvglpnesqalsqrit 664
           |:| ||||||| : |||  |:|: .   .  || |. |   | .  | .. | . . |.
Qy     146 KSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQRSADE-QL-HCAGKAKRSGGHAPRVR 203

Db     665 fspgqdiqlvp-plinllmsiepdviyaghdntkpdtsssllts lnqlgerqllsvvkws 723
                :  | |  |.  |.  ||  .  . .. .. | | .|.. || |....|. .. |.
Qy     204 -ELLLDA-LSPEQLVLTLLEAEPPHVLISRPSA-PFTEASMMMSLTKLADKELVHMISWA 260

Db     724 kslpgfrnlhiddqitliqyswmslmvfglgwrsykhvsgqmlyfapdliln-eq-r-mk 780
           | :||| :| : ||: |::  || ::: || |||  | :| :: ||||:|: :: : :
Qy     261 KKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDH-PGKLI-FAPDLVLDRDEGKCVE 318

Db     781 e lsfyslcltmwqipqefvklqvtheeflcmkvllllnt l-pleglrsqsqfeemrss 837
            | :: : | :        |  | : | |:||:| ::|||. . ||       ..:  : :
Qy     319 GILEIFDM-L-L-ATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAH 375

Db     838 yirelikaig-lrqkgvvptsqrfyqltklldslhdlvkq 876
            :  : |:  :  |. ... |.  .|..|| |   |..
Qy     376 LLNAVTDALVWVIAKSGISSQQQSMRLANLL-MLLSHVRH 414
```

```
RESULT   12
ID    PRGR_HUMAN     STANDARD;      PRT;   933 AA.
AC    P06401;
DT    01-JAN-1988 (REL. 06, CREATED)
DT    01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT    01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE    PROGESTERONE RECEPTOR (PR).
GN    PGR.
OS    HOMO SAPIENS (HUMAN).
OC    EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC    EUTHERIA; PRIMATES.
RN    [1]
RP    SEQUENCE FROM N.A.
RX    MEDLINE; 90228361.
RA    KASTNER P., KRUST A., TURCOTTE B., STROPP U., TORA L., GRONEMEYER H.,
RA    CHAMBON P.;
RL    EMBO J. 9:1603-1614(1990).
RN    [2]
RP    SEQUENCE FROM N.A.
RX    MEDLINE; 87184565.
RA    MISRAHI M., ATGER M., D'AURIOL L., LOOSFELT H., MERIEL C.,
RA    FRIDLANSKY F., GUIOCHON-MANTEL A., GALIBERT F., MILGROM E.;
RL    BIOCHEM. BIOPHYS. RES. COMMUN. 143:740-748(1987).
CC    -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC        THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC        PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC    -!- SUBCELLULAR LOCATION: NUCLEAR.
CC    -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC        A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC    -!- ALTERNATIVE PRODUCTS: THERE ARE TWO FORMS OF THE PROGESTERONE
CC        RECEPTOR: FORM A (THAT STARTS WITH MET-1) AND FORM B (THAT STARTS
CC        WITH MET-165).
CC    -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC        NUCLEAR HORMONE RECEPTORS.
DR    EMBL; X51730; G35652; -.
DR    EMBL; M15716; G189935; -.
DR    PIR; A03245; QRHUP.
DR    PIR; S09971; S09971.
DR    HSSP; P06536; 1GDC.
DR    MIM; 264080; -.
DR    PROSITE; PS00031; STEROID_FINGER.
DR    TRANSFAC; T00696; -.
KW    RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW    ZINC-FINGER; STEROID-BINDING; PHOSPHORYLATION; ALTERNATIVE SPLICING.
FT    DOMAIN        1    566       MODULATING, PRO-RICH.
FT    DNA_BIND    567    632       C4-TYPE ZINC FINGERS (TWO).
FT    ZN_FING     567    587       C4-TYPE.
FT    ZN_FING     603    627       C4-TYPE.
FT    DOMAIN      681    933       STEROID-BINDING.
FT    DOMAIN      183    187       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT    MOD_RES      41     41       PHOSPHORYLATION (BY CK).
FT    MOD_RES     227    227       PHOSPHORYLATION (BY CK).
FT    MOD_RES     232    232       PHOSPHORYLATION (BY CK).
FT    MOD_RES     552    552       PHOSPHORYLATION (BY CK).
FT    MOD_RES     793    793       PHOSPHORYLATION (BY CK).
FT    VARSPLIC      1    164       MISSING (IN FORM B).
FT    CONFLICT    226    226       S -> G (IN REF. 2).
FT    CONFLICT    256    256       C -> V (IN REF. 2).
FT    CONFLICT    344    344       T -> S (IN REF. 2).
FT    CONFLICT    660    660       L -> V (IN REF. 2).
SQ    SEQUENCE   933 AA;  99042 MW;  0D1124EB CRC32;

  Query Match            18.9%;  Score 666;  DB 7;  Length 933;
  Best Local Similarity  33.4%;  Pred. No. 3.81e-124;
  Matches   110;  Conservative   89;  Mismatches 113;  Indels  17;  Gaps  17;

Db     566 iclicqdeasgchygvltcgsckvffkrameggnylcagrndcivdkirrkncpacrlr 625
           :| :|:| ||| |||| :| :|| ||||:::|:::|:|:: |:| :|| |||:| |||||
Qy      95 RCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPAINQCTIDKNRRKSCQACRLR 154

Db     626 kccqagmvl-ggrkfk-kfnkvrvvraldavalpqpvgvpnesqal-sqrftfspgqdiql 683
           || : |||  |:|: :   :  ||  |: |    | :  | :: | : : |        | |
FT   NON_TER       1      1
FT   DOMAIN        1    314       STEROID-BINDING (POTENTIAL)
SQ   SEQUENCE   314 AA;  34872 MW;  0B18C899 CRC32;

  Query Match             30.5%;  Score 1076;  DB 3;  Length 314;
  Best Local Similarity   60.3%;  Pred. No. 2.57e-224;
  Matches  146;  Conservative   45;  Mismatches  46;  Indels   5;  Gaps   5;

Db       1 vlllqgaeppilcsrqkmsrpytevtimtlltsmadkelvhmitwakklpgflqlslhd 60
            :|  ||::    | :|  :|     : |.||  ..|   ||  .||||||·||·||||·|||··|||  |
Qy     216 VLTLLEAEPPHVLISRPS-A-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 273

Db      61 qvlllesswlevlmigliwrsiqcpgklifeedlildrneqtcveqmaeifdmllatv-r 119
           || |||| |:||||:||:||||· ||||||  || |||·|| ||||· ||||||||   :
Qy     274 QVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSR 333

Db     120 frvlklkpeefvclkaiillnsgafsfctgtmeplhdsvavqhmldtitdalifhishfg 179
           || |||    |:::::||:|||||: ::: |:| ::    |   : |:|:::||||:   |:   |
Qy     334 FRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSRKLAHLLNAVTDALVWVIAKSG 392

Db     180 qsqqqsrrqaqllllshirhmsnkgmehlysmkcknkvplydlllemldahrlhrpvk 239
            |:||||  |  |  ||:||||:|| ||||||||  .|||||  ||.||||||||.||  .  |   |
Qy     393 ISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNAH-VLRGCK 451

Db     240 ps 241
           ::
Qy     452 SS 453

RESULT    9
ID   ERR2_HUMAN     STANDARD;      PRT;   433 AA
AC   P11475;
DT   01-OCT-1989 (REL. 12, CREATED)
DT   01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)
DT   01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE   STEROID HORMONE RECEPTOR ERR2.
GN   ERR2.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=HEART;
RX   MEDLINE; 88122546.
RA   GIGUERE V., YANG N., SEGUI P., EVANS R.M.;
RL   NATURE 331:91-94(1988).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; X51417; G36611; -.
DR   PIR; B29345; B29345.
DR   HSSP; P06536; 1GDC.
DR   PROSITE; PS00031; STEROID_FINGER.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER.
FT   DNA_BIND    103    168       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     103    123       C4-TYPE.
FT   ZN_FING     139    163       C4-TYPE.
SQ   SEQUENCE   433 AA;  48287 MW;  673BF88F CRC32;

  Query Match             22.8%;  Score 807;  DB 3;  Length 433;
  Best Local Similarity   39.5%;  Pred. No. 2.95e-159;
  Matches  140;  Conservative   84;  Mismatches 108;  Indels  22;  Gaps  14;

Db     101 rlclvcgdiasgyhygvasceackaffkrtiqgnieyscpatneceitkrrrkscqacrf 160
           ::: ||:| ||·||||| |||:||||||||:|||: :| ||||:| | | ||||||||:
Qy      94 HFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRL 153

Db     161 mkclkvgmlkegvrldrv---r--ggrqky-kirldsenspylslqisppakkpltkivs-- 214
            ||  ||||:| | | :|   |   ||.   :|    : |   :|     |  ·|
Qy     154 RKCYEVGMVKCGSRRERCGYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSPE 213

Db     215 yl-l-vaepdklyam-ppddvpegdikalttlcdladrelvfliswakhipgfsnltlgd 271
```

```
            | | : |::     : :     |   :    : :|   |||:|'| :'|·|· |||· :|:'
Qy     214 EQLVLTLLEAEPPHVLISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFD 273

Db     272 qmsllqsawmeililgivyrslpyddklayaedyimdeehsrlv-gllelyrailqlvrr 330
           |. ||.| |||.|.:|.. ||.     || .| |  ..|  .  :: | |:|||::  :|     |
Qy     274 QVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSR 333

Db     331 ykklkvekeefvmlkalalansdsmy--i---enleavqklqdllh---ealq-dyels- 380
           :: ||::  |:: :||: | || |||   :     :: :: :||  ||:    :||        |
Qy     334 FRELKLQHKEYLCVKAMILLNS-SMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKSG 392

Db     381 -qrheeprragkllltlpllrqtaakavqhfysvklqgkvpmhklflemleakv 433
               ....  | ..||. |. .|... |...|. ::|     ||:  |:||||:| |
Qy     393 ISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNAHV 446

RESULT   10
ID   ERR1_HUMAN     STANDARD;      PRT;   521 AA.
AC   P11474;
DT   01-OCT-1989 (REL. 12, CREATED)
DT   01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)
DT   01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE   STEROID HORMONE RECEPTOR ERR1.
GN   ERR1.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=KIDNEY;
RX   MEDLINE; 88122546.
RA   GIGUERE V., YANG N., SEGUI P., EVANS R.M.;
RL   NATURE 331:91-94(1988).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; X51416; G36609; -.
DR   PIR; A29345; A29345.
DR   HSSP; P06536; 1GDC.
DR   PROSITE; PS00031; STEROID_FINGER.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER.
FT   DNA_BIND    176    241       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     176    196       C4-TYPE.
FT   ZN_FING     212    236       C4-TYPE.
SQ   SEQUENCE   521 AA;  55742 MW;  E2D723EA CRC32;

  Query Match             22.7%;  Score 800;  DB 3;  Length 521;
  Best Local Similarity   39.9%;  Pred. No. 1.49e-156;
  Matches  141;  Conservative   74;  Mismatches 128;  Indels  10;  Gaps   7;

Db     174 rlclvcgdvasgyhygvasceackaffkrtiqgsieyscpasneceitkrrrkacqacrf 233
           ::| ||:| ||||||| |||:||||||||:|||  :| |||:|:| | | |||:|||||:
Qy      94 HFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRL 153

Db     234 tkclrvgmlkegvrldrvrggrqkykrrpevdplpfpg-p-fpagplavaggprktaapv 291
            ||  ||:|| | | :| | |     ::|:  : |  :| :  ::|   :          | :
Qy     154 RKCYEVGMVKCGSRRERC-GYRLVRRQRSADEQLHCAGKAKRSGGHAPRVRELLLDALSP 212

Db     292 nalvshllvvepeklyampdpagpdghlpavatlcdlfdrelvvtiswaksipgfsslsl 351
           : ||  ||  ||  :  :: |::|     : : :|  | |:|:|  ||||| ||||  |||
Qy     213 EQLVLTLLEAEPPHVL-ISRPSAPFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSL 271

Db     352 sdqmsvlqsvwmevlvlgvaqrslplqdelafaedlvld-eegaraaglgelgaallqlv 410
            ||: :|:| ||||||::|:  ||:      | || ||||| :||      |: |:    '|
Qy     272 FDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATT 331

Db     411 rrlqalrlereeyvllkalalansdsvhiedeprlwsscekllh-eal---leyeagra 465
            |.. |.|.. ||. :||: | ||      :     :     |   || | :|:   |    :::
Qy     332 SPFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADSSRKLAHLLNAVTDALVWVIAKS 391

Db     466 gpgggaerrragrllltlpllrqtagkvlahfygvklegkvpmhklflemlea 518
           | ::  :  | : ||: |: :|::: | : |:  :|     ||:  ::|!||:|
DR   HSSP; P06536; 1GDC.
DR   PROSITE; PS00031; STEROID_FINGER.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
FT   DOMAIN        1    146       MODULATING (POTENTIAL).
FT   DNA_BIND    147    212       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     147    167       C4-TYPE.
FT   ZN_FING     183    207       C4-TYPE.
FT   DOMAIN      213    574       STEROID-BINDING (POTENTIAL).
SQ   SEQUENCE   574 AA;  62872 MW;  B391AD02 CRC32;

  Query Match             42.9%;  Score 1516;  DB 3;  Length 574;
  Best Local Similarity   56.3%;  Pred. No. 0.00e+00;
  Matches  225;  Conservative   79;  Mismatches   75;  Indels   22;  Gaps   17;

Db     143 etrycavcsdfasgyhygvwscegckaffkrsiqghndymcpatnqctmdrnrrkscqac 202
           ::::|||||||:|||||||||||||||||||||||||||||:||||||||:|:|||||||||
Qy      92 DAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQAC 151

Db     203 rlrkcyevgmvkgglrkdrgg-rvlrkdkryqgpagdrekptvtwstgqrpgdsgrnsss 261
           |||||||||||| | |::| | |::|:: |    :| | |      : : :|: ||: ::
Qy     152 RLRKCYEVGMVKCGSRRERCGYRLVRRQ-R---SA-D-EQLHCA-GKAKRS--GGH--AP 200

Db     262 slrggqwrgpritmppeqvlfllqggtp-alesrqkvarpytevtmmtllltsmadkelv 320
             :         : :: | : || ||::: |   | ||    | |:|| :||   || :||||||
Qy     201 RVRELL--LDA-LS-PEQLVLTLLEAEPPHVLISRPS-A-PFTEASMMMSLTKLADKELV 254

Db     321 hmiawakkvpgfqelslhdqvqllesswlevlmigliwrsihcpgklifaqdlildrseg 380
           |||:|||||:|||| |||| |||:|||| |:|||||:||:||||  ||||||| ||:||| ||
Qy     255 HMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEG 314

Db     381 dcvegmaeifdmllatvsrfqmlklkpeefvclkaillnpgafsfcsnsveslhnssav 440
            ||||: ||||||||| |||  |||    |::|:||:||||.. ... . .. .| .
Qy     315 KCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSRKL 373

Db     441 esmldnitdalihhishsgasvqqqprrqaqllllshirhmsnkgmehlysikcknkvp 500
              :|: :||||.  |: || | |||. | | ||:||||.|| |||||||| .:||||| ||
Qy     374 AHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVP 433

Db     501 lydlllemldghrlqapgkvaqaqeqteqpstttttstqs 540
           :|||||||||::| |::  | : :| :  :|:  : :  ||
Qy     434 VYDLLLEMLNAHVLRGC-KSSITGSEC-SPAEDSKSKEGS 471

RESULT    7
ID   ESTR_SALSA     STANDARD;      PRT;   535 AA.
AC   P50242;
DT   01-OCT-1996 (REL. 34, CREATED)
     01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
     01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
     ESTROGEN RECEPTOR (ER) (FRAGMENT).
GN   ESR.
OS   SALMO SALAR (ATLANTIC SALMON).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; PISCES; GNATHOSTOMATA;
OC   OSTEICHTHYES; ACTINOPTERYGII; SALMONIFORMES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RA   ROGERS S.A., LLEWELLYN L., RAMSURN V.P., SWEENEY G.E., WIGHAM T.;
RL   SUBMITTED (JUL-1995) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; X89959; G929902; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
FT   NON_TER       1      1
FT   DOMAIN       <1    104       MODULATING (POTENTIAL).
FT   DNA_BIND    105    170       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     105    125       C4-TYPE.
FT   ZN_FING     141    165       C4-TYPE.
FT   DOMAIN      171    535       STEROID-BINDING (POTENTIAL).
SQ   SEQUENCE   535 AA;  58900 MW;  0F71CD0C CRC32;

  Query Match             42.6%;  Score 1503;  DB 3;  Length 535;
  Best Local Similarity   55.7%;  Pred. No. 0.00e+00;
  Matches  224;  Conservative   75;  Mismatches   80;  Indels   23;  Gaps   16;

Db     101 etrycavcsdfasgyhygvwscegckaffkrsiqghndyvcpatnqctmdrnrrkscqac 160
           ::::|||||||:||||||| ||||||||||||||||||||:|||||||||:|:|||||||||
Qy      92 DAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQAC 151

Db     161 rlrkcyevgmvkgglrkdrgg-rvlrkdkryqqpagdrekpyqdlehrtappqdsgrnss 219
           |||||||||||| | |::| | |::|:: |    :| : :    :      : |  : ||: :
Qy     152 RLRKCYEVGMVKCGSRRERCGYRLVRRQ-R----SADE-QLHCAG--K-AKRS-GGHAPR 201

Db     220 ssslsggqwcgpritmppeqvlfllqgaeppalcsrqkvarpytevtmmtllltsmadke 279
               |           : :: | : || ||::  |   | ||    | |:|| :||   || :||||
Qy     202 VRELL----L-DA-LS-PEQLVLTLLEAEPPHVLISRPS-A-PFTEASMMMSLTKLADKE 252

Db     280 lvhmiawakkvpgfqelslhdqvqllesswlevlmigliwrsihcpgkllfaqdllldrn 339
           |||||:|||||:|||| |||| |||:|||| |:|||||:||:||||  ||||||| ||:|||
Qy     253 LVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRD 312

Db     340 eqdcvegmaeifdmllatvsrfrmlklkpeefvclkailllnsgafsfcsnsveslhnss 399
           || ||||: ||||||||| |||| |||    |::|:||:||||: ::: : : ::  :|
Qy     313 EGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDA-DSSR 371

Db     400 avesmldnitdalihhishsgasvqqqprrqvqlilllshirhmsnkgmehlysikcknk 459
             :  :|: :||||:  |: || | |||: |   ||:|||:|| |||||||| ::||||
Qy     372 KLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNV 431

Db     460 vplydlllemldghrlqspgkvaqaqeqteqpstttttstqs 501
           ||:|||||||||::| |::  | : :| :  :|:  : :  ||
Qy     432 VPVYDLLLEMLNAHVLRGC-KSSITGSEC-SPAEDSKSKEGS 471

RESULT    8
ID   ESTR_OREAU     STANDARD;      PRT;   314 AA.
AC   P50240;
DT   01-OCT-1996 (REL. 34, CREATED)
DT   01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   ESTROGEN RECEPTOR (ER) (FRAGMENT).
GN   ESR.
OS   OREOCHROMIS AUREUS.
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; PISCES; GNATHOSTOMATA;
OC   OSTEICHTHYES; ACTINOPTERYGII; PERCIFORMES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BLOOD;
RX   MEDLINE; 96272302.
RA   TAN N., LAM T., DING J.;
RL   DNA SEQ. 5:359-370(1995).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; Z46665; E212276; -.
DR   EMBL; Z46666; E212276; JOINED.
DR   EMBL; Z46667; E212276; JOINED.
DR   EMBL; Z46668; E212276; JOINED.
DR   EMBL; Z46669; E212276; JOINED.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
```

```
Qy     20 TTSPNVLWPTPGHLSP-LVVHRQLSHLYAEPQKSPWCEARSLEHTLPVNRETLKRKVSG- 77

Db    168 erlssssekgnmimesaketrycavcndyasgyhygvwscegckaffkrsiqghndymcp 227
          :|  :|:  |      | ::::||||:|||||||||||||||||||||||||||||||.||
Qy     78 NR-CASPVTGP---GSKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICP 133

Db    228 atnqctidknrrksqacrlrkcyevgmmkggirkdrrggrmlkhkrqrddlegrnemgt 287
          ||||||||||||||||||||||||||:| | |::| | |:::: |  |: :  .    |
Qy    134 ATNQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVRRQRSADE-QLHCA-GK 191

Db    288 sgdmraanlwpsplvikhtkknspalsltadqmvsalldaepplíysqydpsrpfseasm 347
          : :|::    :| | :    :  |||  :.| |  || |||          || ||.||||
Qy    192 A-KRSGG--HAPRV-RELLLD--ALS--PEQLVLTLLEAEPPHVLIS-RPSAPFTEASM 241

Db    349 mqlltnladrelvhminwakrvpgfgdlnlhdqvhllecawleilmiglvwrsmehpgkl 407
          |  .|:||||:|||||:|||::||  .|.| |||.|||  |:|:||:||:|||::|||||
Qy    242 MMSLTKLADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKL 301

Db    408 lfapnllldrnqgkcvegmveifdmllatssrfrmmnlqgeefvclksiillnsgvytfl 467
          :|||:|:|||::||||||::|||||||||:||||  ::||  |:.|.|..|||||..| ::
Qy    302 IFAPDLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLV 361

Db    468 sstlksleekdhihrvldkindtlihlmakagltlqqqhrrlaqlllilshirhmsnkgm 527
          ::|  : :     : ::|: : |:|: ::||.|..  |||  ||| ||..|||:|| |||||
Qy    362 TATQDA-DSSRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGM 420

Db    528 ehlynmkcknvvplydlllemldahrlhapasrmgvppeepsqsqltttsstsahs 583
          ||| ||||||||:|||||||||:|| |::  |  :  :  |::   :  :| :::|
Qy    421 EHLLNMKCKNVVPVYDLLLEMLNAHVLRGCKSSITGSECSPAEDSKSKEGSQNPQS 476
```

```
RESULT     5
ID   ESTR_ORYLA      STANDARD;       PRT;   575 AA.
AC   P50241;
DT   01-OCT-1996 (REL. 34, CREATED)
DT   01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   ESTROGEN RECEPTOR (ER).
GN   ESR.
OS   ORYZIAS LATIPES (MEDAKA FISH).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; PISCES; GNATHOSTOMATA;
OC   OSTEICHTHYES; ACTINOPTERYGII; ATHERINIFORMES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RA   OKADA H.;
RL   SUBMITTED (APR-1994) TO EMBL/GENBANK/DDBJ DATA BANKS
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS
DR   EMBL; D28954; G602326; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
FT   DOMAIN        1    140       MODULATING.
FT   DNA_BIND    141    206       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     141    161       C4-TYPE.
FT   ZN_FING     177    201       C4-TYPE.
FT   DOMAIN      207    575       STEROID-BINDING.
FT   DOMAIN      254    262       POLY-GLY.
FT   DOMAIN      275    278       POLY-LEU.
FT   DOMAIN      466    470       POLY-LEU.
FT   DOMAIN      531    534       POLY-GLY.
SQ   SEQUENCE   575 AA,  62987 MW.  D9FE40C0 CF732.

  Query Match              44.7%,  Score 1578;  DB 3;  Length 575;
  Best Local Similarity  53.9%;  Pred. No. 0.00e-00;
  Matches  249;  Conservative   97;  Mismatches 103;  Indels  23;  Gaps  16;
```

```
Db     52 ppsegslqs-lqsgptsplvfvpssprlspfm-hpp-shhylettstpvyrsshqaasre 108
          |::  :|::  |   ||| |: |:: :|||::  |    || | |     :|    :       :
Qy      6 PSNVTNLEGGPGRQTTSPNVLWPTPGHLSPLVVHRQLSHLYAEPQKSPWCEARSLEHTLP 65

Db    109 dqcgsredtcslgelgagagaggfemakdtrfcavcsdyasgyhygvwscegckaffkrs 168
              .     | |.  |.. .|     . .|..||||||||||:|||||||||||||||||
Qy     66 VNRETLKRKVS-GNRCASPVTGPGSK-RDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRS 123

Db    169 iqghndymcpatnqctidrnrrksсqacrlrkcyevgmmkggvrkdrgrvlrrdkrrtgv 228
          |||||||.|||||||||||.|||||||||||||||||||:| | |::|      '   ||
Qy    124 IQGHNDYICPATNQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGY-RLV-RR--- 178

Db    229 gdgdvvkggekktvhygrkrsstgggggggglsvtslppeqvllllqgaepp-llr 287
             ..    ::|:  :|  |: : | ||  :        | : :::|||::| |  |||| :|
Qy    179 ---QR--SADEQ--LHCAGKAKRS-GGHAPRVRELL-LDALSPEQLVLTLLEAEPPHVLI 229

Db    288 srqklsrpytevtmmtlltsmadkelvhmiawakklpgflqlslhdqvlllesswlevlm 347
          ||    : |:|| :||   || :||||||||||:|||.:||||:|||| |||  |||'  :|||
Qy    230 SRPS-A-PFTEASMMMSLTKLADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLM 287

Db    348 igliwrsihcpgklifaqdlildrnegdcvegmteifdmllatasrfrvlklkpeefvcl 407
          :||:||||  |||||||  ||:|||:||  ||||: |||||||||:|||:| |||    |::|:
Qy    288 MGLMWRSIDHPGKLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCV 347

Db    408 kaiillnsgafsfctgtmeplhnsaavqsmldtitdalihylsqsgylaqeqarrqaqll 467
          ||.|||||. ... |.| ..   .|   .   .|::|||||.  |. ||  :|:|: | | ||
Qy    348 KAMILLNSSMYPLVTATQDA-DSSRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLL 406

Db    468 lllshirhmsnkgmehlysmkcknkvplydllleml dahrlh 509
          :||||:|| ||||||||| :||||| ||:||||||||:|| |:
Qy    407 MLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNAHVLR 448
```

```
RESULT     6
ID   ESTR_SALIR      STANDARD;       PRT;   574 AA.
AC   P16058;
DT   01-APR-1990 (REL. 14, CREATED)
DT   01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   ESTROGEN RECEPTOR (ER).
GN   ESR.
OS   SALMO IRIDEUS (RAINBOW TROUT).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; PISCES; GNATHOSTOMATA;
OC   OSTEICHTHYES; ACTINOPTERYGII; SALMONIFORMES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 91006824.
RA   PAKDEL F., LE GAC F., LE GOFF P., VALOTAIRE Y.;
RL   MOL. CELL. ENDOCRINOL. 71:195-204(1990).
RN   [2]
RP   SEQUENCE OF 150-574 FROM N.A.
RC   TISSUE=LIVER;
RX   MEDLINE; 89127284.
RA   PAKDEL F., LE GUELLEC C., VAILLANT C., LE ROUX M.G., VALOTAIRE Y.;
RL   MOL. ENDOCRINOL. 3:44-51(1989).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- IN THE ABSENCE OF LIGAND, STEROID HORMONE RECEPTORS ARE THOUGHT
CC       TO BE WEAKLY ASSOCIATED WITH NUCLEAR COMPONENTS. HORMONE BINDING
CC       GREATLY INCREASES RECEPTOR AFFINITY. THE HORMONE-RECEPTOR
CC       COMPLEX APPEARS TO RECOGNIZE DISCRETE DNA SEQUENCES UPSTREAM OF
CC       TRANSCRIPTIONAL START SITES.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; M31559; G213784; -.
DR   PIR; A37197; A37197.
DR   PIR; A4[illegible]; A4[illegible].
```

```
RESULT     3
ID   ESTR_MOUSE      STANDARD;      PRT;   599 AA.
AC   P19785;
DT   01-FEB-1991 (REL. 17, CREATED)
DT   01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
DT   01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE   ESTROGEN RECEPTOR (ER).
GN   ESR OR ESTR.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=UTERUS;
RX   MEDLINE; 91042558.
RA   WHITE R., LEES J.A., NEEDHAM M., HAM J., PARKER M.;
RL   MOL. ENDOCRINOL. 1:735-744(1987).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- IN THE ABSENCE OF LIGAND, STEROID HORMONE RECEPTORS ARE THOUGHT
CC       TO BE WEAKLY ASSOCIATED WITH NUCLEAR COMPONENTS; HORMONE BINDING
CC       GREATLY INCREASES RECEPTOR AFFINITY. THE HORMONE-RECEPTOR
CC       COMPLEX APPEARS TO RECOGNIZE DISCRETE DNA SEQUENCES UPSTREAM OF
CC       TRANSCRIPTIONAL START SITES.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; M38651; G193180; -.
DR   PIR; A40061; A40061.
DR   HSSP; P06536; 1GDC.
DR   PROSITE; PS00031; STEROID_FINGER.
DR   TRANSFAC; T00259; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
FT   DOMAIN        1    188       MODULATING.
FT   DNA_BIND    189    254       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     189    209       C4-TYPE.
FT   ZN_FING     225    249       C4-TYPE.
FT   DOMAIN      255    599       STEROID-BINDING.
FT   DOMAIN       64     72       POLY-ALA.
SQ   SEQUENCE   599 AA;  66955 MW;  6FD9EADB CRC32;

  Query Match             47.3%;  Score 1669;  DB 3;  Length 599;
  Best Local Similarity   51.7%;  Pred. No. 0.00e+00;
  Matches  246;  Conservative  135;  Mismatches 164;  Indels  27;  Gaps  14;

Db     108 spsplmllhpppqlspflhphgqqvpyylenepsay-avrdtgppafyrsnsdnrrqngr 166
            : || :|     ' :||| |   | |       | | : |:       |              :: :|: :|
Qy      20 TTSPNVLWPTPGHLSP-LVVHRQLSHIYAEPQKSPWCEARSLEHTLPVNRETLKRKVSG- 77

Db     167 erlsssnekqnmimesaketrycavcndyasgyhygvwscegckaffkrsiqghndymcp 226
            :|   :|     |        | :::::||||:||||||||||||||||||||||||||||:||
Qy      78 NR-CASPVTGP---GSKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICP 133

Db     227 atnqctidknrrksccacrlrkcyevgmmkggirkdrrggrmlkhkrqrddlegrnemga 286
            |||||||||||||||||||||||||||:|  | |::| | |::: |||    : :     |
Qy     134 ATNQCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRLVR--RQRSA-DEQLH-CA 190

Db     287 sgdmraanlwpsplvikhtkknspalsltadqmvsalldaeppmiysqydpsrpfseasm 346
            :  :|::      :| | :      :  |||  ::|:| :||:|||| :          || ||:||||
Qy     190 GKAKRSGG--HAPRV-RELLLD--ALS--PEQLVLTLLEAEPPHVLIS-RPSAPFTEASM 241

Db     347 mqlltnladrelvhminwakrvpgfvdltlhdqvhllecawleilmiglvwrsmehpgkl 406
            |   ||:|||:||||||||:||||::||| :|:| ||||:|||   |:|:|||:||:||||::||||||
Qy     242 MMSLTKLADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKL 301

Db     407 lfapnllldrnqgkcvegmveifdmllatssrfrmmnlqgeefvclksiillnsgvytfl 466
            :|||:|:|||::|||||||:.|||||||||.||||  .:||   |..|.|..!||||..| ..
Qy     302 IFAPDLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYICVKAMILLNSSMYPLV 361

Db     467 sstlksleekdhihrvldkitdtlihlmakagltlqqqhrrlaqllllshirhmsnkgm 526
            ::|  : :      : ::|: :||:|: ::||:|:: |||   ||| ||::|||:|| |||||
Qy     362 TATQDA-DSSRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGM 420

Db     527 ehlynmkcknvvplydlllemldahrlhapasrmgvppeepsqtqlattsstsahs 582
            ||| ||||||||:|||||||||:|| |::  |  :  :  |::   :  :| :::|
Qy     421 EHLLNMKCKNVVPVYDLLLEMLNAHVLRGCKSSITGSECSPAEDSKSKEGSQNPQS 476
```

```
RESULT     4
ID   ESTR_RAT        STANDARD;      PRT;   600 AA.
AC   P06211;
DT   01-JAN-1988 (REL. 06, CREATED)
DT   01-JAN-1988 (REL. 06, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   ESTROGEN RECEPTOR (ER).
GN   ESR OR ESTR.
OS   RATTUS NORVEGICUS (RAT).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WISTAR;
RA   MURAMATSU M.;
RL   SUBMITTED (MAR-1987) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 87174780.
RA   KOIKE S., SAKAI M.;
RL   NUCLEIC ACIDS RES. 15:2499-2513(1987).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=UTERUS;
RA   MAGGI A.M.A.;
RL   SUBMITTED (JUN-1991) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- IN THE ABSENCE OF LIGAND, STEROID HORMONE RECEPTORS ARE THOUGHT
CC       TO BE WEAKLY ASSOCIATED WITH NUCLEAR COMPONENTS; HORMONE BINDING
CC       GREATLY INCREASES RECEPTOR AFFINITY. THE HORMONE-RECEPTOR
CC       COMPLEX APPEARS TO RECOGNIZE DISCRETE DNA SEQUENCES UPSTREAM OF
CC       TRANSCRIPTIONAL START SITES.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; Y00102; G56111; -.
DR   EMBL; X61098; G56121; -.
DR   PIR; S07379; S07379.
DR   HSSP; P06536; 1GDC.
DR   PROSITE; PS00031; STEROID_FINGER.
DR   TRANSFAC; T00258; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
FT   DOMAIN        1    189       MODULATING.
FT   DNA_BIND    190    255       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     190    210       C4-TYPE.
FT   ZN_FING     226    250       C4-TYPE.
FT   DOMAIN      256    600       STEROID-BINDING.
FT   DOMAIN       64     71       POLY-ALA.
FT   CONFLICT    488    488       N -> T (IN REF. 3).
SQ   SEQUENCE   600 AA;  67030 MW;  4CCC59B2 CRC32;

  Query Match             47.1%;  Score 1663;  DB 3;  Length 600;
  Best Local Similarity   51.1%;  Pred. No. 0.00e+00;
  Matches  243;  Conservative  110;  Mismatches 103;  Indels  20;  Gaps  14;

Db     109 spsplmllhppphvspflhphghqvpyylenepsay-avrdtgppafyrsndnrrqnar 167
            : || :|    | |:|| |  | :     | | : |:      |             :: :|: :|
CC       GREATLY INCREASES RECEPTOR AFFINITY. THE HORMONE-RECEPTOR
CC       COMPLEX APPEARS TO RECOGNIZE DISCRETE DNA SEQUENCES UPSTREAM OF
CC       TRANSCRIPTIONAL START SITES.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; X03635; G31234; -.
DR   EMBL; M12674; G182193; -.
DR   PIR; A03214; QRHUE.
DR   HSSP; P06536; 1GDC.
DR   MIM; 133430; -.
DR   PROSITE; PS00031; STEROID_FINGER.
DR   TRANSFAC; T00261; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING; POLYMORPHISM.
FT   DOMAIN        1    184       MODULATING.
FT   DNA_BIND    185    250       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     185    205       C4-TYPE.
FT   ZN_FING     221    245       C4-TYPE.
FT   DOMAIN      251    310       HINGE.
FT   DOMAIN      311    551       STEROID-BINDING.
FT   VARIANT     400    400       G -> V (DESTABILIZES THE RECEPTOR AND
FT                                DECREASES ITS AFFINITY FOR ESTRADIOL AT
FT                                25 DEGREES CELSIUS, BUT NOT AT 4 DEGREES
FT                                CELSIUS).
FT   VARIANT     447    447       C -> A (LOSS OF HORMONE BINDING CAPACITY
FT                                AND TEMPERATURE-SENSITIVE LOSS IN DNA-
FT                                BINDING).
SQ   SEQUENCE   595 AA;  66216 MW;  D6945BB9 CRC32;

  Query Match             47.9%;  Score 1693;  DB 3;  Length 595;
  Best Local Similarity   56.8%;  Pred. No. 0.00e+00;
  Matches  229;  Conservative   98;  Mismatches   59;  Indels   17;  Gaps   13;

Db      179 saketrycavcndyasgyhygvwscegckaffkrsiqghndymcpatnqctidknrrksc 237
            | ::::::|||:||||||||||||||||||||||||||||||:||||||||||||||||||
Qy       89 SKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSC 148

Db      238 qacrlrkcyevgmmkggirkdrrggrmlkhkrqrddgegrgevgsagdmraanlwpsplm 297
            ||||||||||||||:| | |::| | |:::  |||    | :   :: : :|::    :|
Qy      149 QACRLRKCYEVGMVKCGSRRERCGYRIVR--RQRSADE-QLHCAGKA-KRSGG--HAP-R 201

Db      298 ikrskknslalsltadqmvsalldaepp-ilyseydptrpfseasmmglltnladrelvh 356
            : |     |   :|::::|:| ::|:|||| :| |    |: ||:||||||  ||:||||:||||
Qy      202 V-RELL--LD-ALSPEQLVLTLLEAEPPHVLISR--PSAPFTEASMMMSLTKLADKELVH 255

Db      357 minwakrvpgfvdltlhdqvhllecawleilmiglvwrsmehpgkllfapnllldrnqgk 416
            ||:|||::||||:|:| |||:|||  |.|.||.||.|||..||||:|||:|:||||::||
Qy      256 MISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFAPDLVLDRDEGK 315

Db      417 cvegmveifdmllatssrfrmmnlqgeefvclksiillnsgvytflsstlksleekdhih 476
            ||||::|||||||||||:|||| ::||  |::|:|::||||||::| ::::|  :  :    :
Qy      316 CVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVIATQDA-DSSRKLA 374

Db      477 rvldkitdtlihlmakagltlqqqhqrlaqlllilshirhmsnkgmehlysmkcknvvpl 536
            ::|: :||:|: ::||:|:: |||  ||| ||::|||:|| |||||||| :|||||||||:
Qy      375 HLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPV 434

Db      537 ydlllemldahrlha-ptsrggasveetdqshlatagstsshs 578
             |||||||:||  ::  :|  |:     :::|  :  || :::|
Qy      435 YDLLLEMLNAHVLRGCKSSITGSECSPAEDSK-SKEGSQNPQS 476

RESULT    2
ID   ESTR_CHICK     STANDARD;      PRT;   589 AA.
AC   P06212;
DT   01-JAN-1988 (REL. 06, CREATED)
DT   01-JAN-1988 (REL. 06, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   ESTROGEN RECEPTOR (ER).
GN   ESR.
OS   GALLUS GALLUS (CHICKEN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;
OC   GALLIFORMES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 86247578.
RA   KRUST A., GREEN S., ARGOS P., KUMAR V., WALTER P., BORNERT J.-M.,
RA   CHAMBON P.;
RL   EMBO J. 5:891-897(1986).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- IN THE ABSENCE OF LIGAND, STEROID HORMONE RECEPTORS ARE THOUGHT
CC       TO BE WEAKLY ASSOCIATED WITH NUCLEAR COMPONENTS; HORMONE BINDING
CC       GREATLY INCREASES RECEPTOR AFFINITY. THE HORMONE-RECEPTOR
CC       COMPLEX APPEARS TO RECOGNIZE DISCRETE DNA SEQUENCES UPSTREAM OF
CC       TRANSCRIPTIONAL START SITES.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC       NUCLEAR HORMONE RECEPTORS.
DR   EMBL; X03805; G63380; -.
DR   PIR; S07192; S07192.
DR   HSSP; P06536; 1GDC.
DR   PROSITE; PS00031; STEROID_FINGER.
DR   TRANSFAC; T00264; -.
KW   RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW   ZINC-FINGER; STEROID-BINDING.
FT   DOMAIN        1    178       MODULATING.
FT   DNA_BIND    179    244       C4-TYPE ZINC FINGERS (TWO).
FT   ZN_FING     179    199       C4-TYPE.
FT   ZN_FING     215    239       C4-TYPE.
FT   DOMAIN      245    304       HINGE.
FT   DOMAIN      305    589       STEROID-BINDING.
SQ   SEQUENCE   589 AA;  66746 MW;  85FF1F4A CRC32;

  Query Match             47.7%;  Score 1685;  DB 3;  Length 589;
  Best Local Similarity   54.0%;  Pred. No. 0.00e+00;
  Matches  241;  Conservative   99;  Mismatches   86;  Indels   16;  Gaps   11;

Db      100 spvvlqtapqlspfihhsqqvpyylenepqgsiqmreaappalyrpasdnrrhsirerm 159
            || |:  |: :||||:: |: |     | | : ::      :    ::    : :  :: :
Qy       22 SPNVLWPTSGHLSPLVVHR-QLSHLYAEPQKSPWCEARSLEHTL-PVNRETLKRKVSGNR 79

Db      160 sstnekgslsmestketrycavcndyasgyhygvwscegckaffkrsiqghndymcpatn 219
             ::    |: :   | :::::|||:|||||||||||||||||||||||||||||:||:|||
Qy       80 CASPVTSP-S--SKRDAHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHNDYICPATN 136

Db      220 qctidknrrkscqacrlrkcyevgmmkggirkdrrggrmmkhkrqrevqldrngeasste 279
            |||||||||||||||||||||||||:| | |::| |   :::: |   :||     | |   :
Qy      137 QCTIDKNRRKSCQACRLRKCYEVGMVKCGSRRERCGYRIVRRQRSADEQLHCAGKAKRSG 196

Db      280 lraptlwtsplvvhhnkhnspalsltadqmvsalleaepplvysqydpnrpfneasmmtl 339
             :|| ::   |::        :  |||  :||:| :||||||| |        |: || ||||:
Qy      197 GHAPRVR-E-LLL-----D--ALS--PEQLVLTLLEAEPPHVLIS-RPSAPFTEASMMMS 244

Db      340 ltnladrelvhminwakrvpgfvdltlhdqvhllecawleilmiglvwrsmehpgkllfa 399
            ||:|||:||||||||:||||::||||:|:| |||:|||  |:|:||:||:||||::  ||||:||
Qy      245 LTKLADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFA 304

Db      400 pnllldrnqgkcvegmveifdmllatsarfrmmnlqgeefvclksiillnsgvytflsst 459
            |:|:|||::|||||||::|||||||||||::||| ::||  |::|:|::||||||::| ::::|
Qy      305 PDLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVIAT 364

Db      460 lksleekdhihrvldkitdtlihlmaksglslqqqhrrlaqlllilshirhmsnkgmehl 519
              :    |    | :|: :||:|: ::||:|:: |||  ||| ||::|||:|| |||||.||
Qy      365 QDADSSRKLAH-LLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGMEHL 423

Db      520 ynmkcknvvplydlllemldahrlha 545
             |||||||||:||||||||:|| |::
Qy      424 LNMKCKNVVPVYDLLLEMLNAHVLRG 449
```

```
***********************************************************************
 ___   ___   _______   _______   _______   _______   _     _
|   \ /   | |   __  | |   ____| |   __  | |   ____| | |   | |
|    V    | |  |__| | |  |____  |  |__| | |  |      | |___| |
|  |\ /|  | |   ____| |____   | |   _  _| |  |      |  ___  |
|  | V |  | |  |       ____|  | |  | \ \  |  |____  | |   | |
|__|   |__| |__|      |_______| |__|  \_\ |_______| |_|   |_|  (TM)

***********************************************************************

        Release 2 1D John F  Collins, Biocomputing Research Unit
        Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
              Distribution rights by IntelliGenetics, Inc.

MPsrch_pp   protein - protein database search, using Smith-Waterman algorithm

Run on:          Thu Dec 18 08:44:58 1997;  MasPar time 12.29 Seconds
                                            822.984 Million cell updates/sec
ular output not generated

Title:           >US-08-826-361A-5
Description:     (1-477) from US08826361A.pep
Perfect Score:   3532
Sequence:        1 MNYSIPSNVTNLEGGPGRQT         ECSPAEDSKSKEGSQNPQSQ 477

Scoring table:   PAM 150
                 Gap 11

Searched:        59021 seqs, 21210388 residues

Post-processing. Minimum Match 0%
                 Listing first 45 summaries

Database:        swiss-prot34
                   1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
                   8:part8 9:part9 10:part10 11:part11

Statistics:      Mean 50.037,  Variance 83.321,  scale 0.596

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                SUMMARIES
                 %
ult            Query
No.   Score    Match Length DB  ID            Description             Pred. No.
-------------------------------------------------------------------------------
  1    1693    47.9    595   3  ESTR_HUMAN    ESTROGEN RECEPTOR (ER   0.00e+00
  2    1685    47.7    589   3  ESTR_CHICK    ESTROGEN RECEPTOR (ER   0.00e+00
  3    1669    47.3    599   3  ESTR_MOUSE    ESTROGEN RECEPTOR (ER   0.00e+00
  4    1663    47.1    600   3  ESTR_RAT      ESTROGEN RECEPTOR (ER   0.00e+00
  5    1578    44.7    575   3  ESTR_ORYLA    ESTROGEN RECEPTOR (ER   0.00e+00
  6    1516    42.9    574   3  ESTR_SALIR    ESTROGEN RECEPTOR (ER   0.00e+00
  7    1503    42.6    535   3  ESTR_SALSA    ESTROGEN RECEPTOR (EP   0.00e+00
  8    1076    30.5    314   3  ESTR_ONCAU    ESTROGEN RECEPTOR (ER   2.57e-224
  9     807    22.9    433   3  ERR2_HUMAN    STEROID HORMONE RECEP   2.95e-158
 10     800    22.7    521   3  ERR1_HUMAN    STEROID HORMONE RECEP   1.49e-156
 11     681    19.3    923   7  PRGR_MOUSE    PROGESTERONE RECEPTOR   9.45e-128
 12     666    18.9    933   7  PRGR_HUMAN    PROGESTERONE RECEPTOR   3.81e-124
 13     662    18.7    930   7  PRGR_RABIT    PROGESTERONE RECEPTOR   3.48e-123
 14     614    17.4    777   4  GCR_HUMAN     GLUCOCORTICOID RECEPT   1.07e-111
 15     608    17.2    776   4  GCR_XENLA     GLUCOCORTICOID RECEPT   2.40e-110
 16     597    16.9    786   7  PRGR_CHICK    PROGESTERONE RECEPTOR   1.22e-107
 17     586    16.6    783   4  GCR_MOUSE     GLUCOCORTICOID RECEPT   5.00e-105
 18     579    16.4    133   3  ESTR_BOVIN    ESTROGEN RECEPTOR (ER   2.30e-103
 19     577    16.3    771   4  GCR_CAVPO     GLUCOCORTICOID RECEPT   6.86e-103
 20     577    16.3    795   4  GCR_RAT       GLUCOCORTICOID RECEPT   6.86e-103
 21     571    16.2    899   1  ANDR_MOUSE    ANDROGEN RECEPTOR.      1.81e-101
 22     565    16.1    909   1  ANDR_RABIT    ANDROGEN RECEPTOR (ER   3.34e-101
```

```
 23     570    16.1    902   1  ANDR_RAT      ANDROGEN RECEPTOR.      3.14e-101
 24     567    16.1    919   1  ANDR_HUMAN    ANDROGEN RECEPTOR.      1.61e-100
 25     562    15.9    981   6  MCR_RAT       MINERALOCORTICOID REC   2.46e-99
 26     544    15.4    758   4  GCR_ONCMY     GLUCOCORTICOID RECEPT   4.43e-95
 27     533    15.1    533   8  RRXB_HUMAN    RETINOIC ACID RECEPTO   1.74e-92
 28     529    15.0    458   8  RRXB_RAT      RETINOIC ACID RECEPTO   1.52e-91
 29     529    15.0    520   8  RRXB_MOUSE    RETINOIC ACID RECEPTO   1.52e-91
 30     523    14.8    984   6  MCR_HUMAN     MINERALOCORTICOID REC   3.92e-90
 31     517    14.6    111   3  ESTR_SHEEP    ESTROGEN RECEPTOR (ER   1.01e-88
 32     514    14.6    467   8  RRXA_MOUSE    RETINOIC ACID RECEPTO   5.09e-88
 33     514    14.6    467   8  RRXA_RAT      RETINOIC ACID RECEPTO   5.09e-88
 34     513    14.5    462   8  RRXA_HUMAN    RETINOIC ACID RECEPTO   8.74e-88
 35     507    14.4    488   8  RRXA_XENLA    RETINOIC ACID RECEPTO   2.23e-86
 36     502    14.2    597   6  NGFI_RAT      NERVE GROWTH FACTOR I   3.31e-85
 37     503    14.2    598   6  NGFI_CANFA    ORPHAN NUCLEAR RECEPT   1.93e-85
 38     501    14.2    598   6  NGFI_HUMAN    EARLY RESPONSE PROTEI   5.67e-85
 39     499    14.1    121   3  ESTR_MACMU    ESTROGEN RECEPTOR (ER   1.66e-84
 40     493    14.0    467   8  RRXA_CHICK    RETINOIC ACID RECEPTO   4.21e-83
 41     491    13.9    601   6  NGFI_MOUSE    NUCLEAR HORMONE RECEP   1.23e-82
 42     489    13.8    463   8  RRXG_MOUSE    RETINOIC ACID RECEPTO   2.52e-82
 43     488    13.8    463   8  RRXG_HUMAN    RETINOIC ACID RECEPTO   6.19e-82
 44     474    13.4    577   6  NGFI_XENLA    NERVE GROWTH FACTOR I   1.13e-78
 45     473    13.4    598   8  RNR1_RAT      REGENERATING LIVER NU   1.93e-78

                                ALIGNMENTS

RESULT     1
ID   ESTR_HUMAN     STANDARD;      PRT;   595 AA.
AC   P03372;
DT   21-JUL-1986 (REL. 01, CREATED)
DT   01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DT   01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE   ESTROGEN RECEPTOR (ER).
GN   ESR.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 86122927.
RA   GREENE G.L., GILNA P., WATERFIELD M., BAKER A., HORT Y., SHINE J.;
RL   SCIENCE 231:1150-1154(1986).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 86146892.
RA   GREEN S., WALTER P., KUMAR V., KRUST A., BORNERT J.-M., ARGOS P.,
RA   CHAMBON P.;
RL   NATURE 320:134-139(1986).
RN   [3]
RP   STRUCTURE BY NMR OF 180-262.
RX   MEDLINE; 91061891.
RA   SCHWABE J.W.F., NEUHAUS D., RHODES D.;
RL   NATURE 348:458-461(1990).
RN   [4]
RP   VARIANT VAL-400.
RX   MEDLINE; 90005402.
RA   TORA L., MULLICK A., METZGER D., PONGLIKITMONGKOL M., PARK I.,
RA   CHAMBON P.;
RL   EMBO J. 8:1981-1986(1989).
RN   [5]
RP   VARIANT ALA-447.
RX   MEDLINE; 92250650.
RA   REESE J.C., KATZENELLENBOGEN B.S.;
RL   J. BIOL. CHEM. 267:9868-9873(1992).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- IN THE ABSENCE OF LIGAND, STEROID HORMONE RECEPTORS ARE THOUGHT
CC       TO BE WEAKLY ASSOCIATED WITH NUCLEAR COMPONENTS; HORMONE BINDING
```

US 0882636104P1

Creation date: 09-04-2003
Indexing Officer: NTRAN2 - NGHI TRAN
Team: OIPEBackFileIndexing
Dossier: 08826361

Legal Date: 04-09-2002

No.	Doccode	Number of pages
1	SRNT	76

Total number of pages: 76

Remarks:

Order of re-scan issued on ....